PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/82, 16/30, C12N 5/08, 15/62, 15/63, C12Q 1/68, A61K 31/7088, 38/17, 39/395, G01N 33/574, A61P 35/00

A2

(11) International Publication Number: **WO 00/52165**

(43) International Publication Date: 8 September 2000 (08.09.00)

(21) International Application Number: PCT/US00/05431

(22) International Filing Date: 29 February 2000 (29.02.00)

(30) Priority Data:

09/262,505	4 March 1999 (04.03.99)	US
09/272,886	19 March 1999 (19.03.99)	US
09/396,313	17 September 1999 (17.09.99)	US

(71) Applicant *(for all designated States except US)*: CORIXA CORPORATION [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: LODES, Michael, J. [US/US]; 9223 – 36th Avenue Southwest, Seattle, WA 98126 (US).

(74) Agents: MAKI, David, J. et al.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104–7092 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: COMPOSITIONS AND METHODS FOR BREAST CANCER THERAPY AND DIAGNOSIS

(57) Abstract

Compositions and methods for the therapy and diagnosis of cancer, such as breast cancer, are disclosed. Compositions may comprise one or more breast tumor antigens, immunogenic portions thereof or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a breast tumor antigen, or a T cell that is specific for cells expressing a breast tumor antigen. Such compositions may be used, for example, for the prevention and treatment of diseases such as breast cancer. Diagnostic methods based on detecting a breast tumor antigen in a sample are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

COMPOSITIONS AND METHODS FOR BREAST CANCER THERAPY AND DIAGNOSIS

TECHNICAL FIELD

The present invention relates generally to therapy and diagnosis of breast cancer. The invention is more specifically related to polypeptides comprising at least a portion of a breast tumor antigen, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for treatment of breast cancer, and for the diagnosis and monitoring of such cancer.

BACKGROUND OF THE INVENTION

Breast cancer is a significant health problem for women in the United States and throughout the world. Although advances have been made in detection and treatment of the disease, breast cancer remains the second leading cause of cancer-related deaths in women, affecting more than 180,000 women in the United States each year. For women in North America, the life-time odds of getting breast cancer are now one in eight.

No vaccine or other universally successful method for the prevention or treatment of breast cancer is currently available. Management of the disease currently relies on a combination of early diagnosis (through routine breast screening procedures) and aggressive treatment, which may include one or more of a variety of treatments such as surgery, radiotherapy, chemotherapy and hormone therapy. The course of treatment for a particular breast cancer is often selected based on a variety of prognostic parameters, including an analysis of specific tumor markers. *See, e.g.,* Porter-Jordan and Lippman, *Breast Cancer 8*:73-100, 1994. However, the use of established markers often leads to a result that is difficult to interpret, and the high mortality observed in breast cancer patients indicates that improvements are needed in the treatment, diagnosis and prevention of the disease.

Immunotherapies have the potential to substantially improve breast cancer treatment and survival. Such therapies may involve the generation or enhancement of an immune response to a breast tumor antigen. However, to date, relatively few breast tumor antigens are known and the generation of an immune response against such antigens has not been shown to be therapeutically beneficial.

Accordingly, there is a need in the art for improved methods for identifying breast tumor antigens and for using such antigens in the diagnosis and therapy of breast cancer. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

Briefly stated, this invention provides compositions and methods for the therapy of cancer, such as breast cancer. In one aspect, the present invention provides polypeptides comprising an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments, the polypeptide comprises a sequence that is encoded by a polynucleotide sequence selected from the group consisting of sequences recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186, and complements of such polynucleotides.

The present invention further provides polynucleotides that encode a polypeptide as described above or a portion thereof, expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions and vaccines. Within one such aspect, a pharmaceutical composition comprises: (a) a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence

selected from the group consisting of: (i) polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and (ii) complements of the foregoing polynucleotides; and (b) a physiologically acceptable carrier.

Within a related aspect, a vaccine comprises: (a) a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (i) polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and (ii) complements of the foregoing polynucleotides; and (b) a non-specific immune response enhancer.

Within further aspects, pharmaceutical compositions provided herein comprise: (a) a polynucleotide encoding at least 15 amino acid residues of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (i) polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and (ii) complements of the foregoing polynucleotides; and (b) a physiologically acceptable carrier.

Within related aspects, the present invention provides vaccines comprising: (a) a polynucleotide encoding at least 15 amino acid residues of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (i) polynucleotides recited in Figures 1A-1L

(SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and (ii) complements of the foregoing polynucleotides; and (b) a non-specific immune response enhancer.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a breast tumor antigen, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (i) polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and (ii) complements of the foregoing polynucleotides; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (i) polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and (ii) complements of the foregoing polynucleotides; and (b) a pharmaceutically acceptable carrier or excipient. Antigen presenting cells include dendritic cells and macrophages.

Within related aspects, vaccines are provided, comprising: (a) an antigen presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (i) polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186;

and (ii) complements of the foregoing polynucleotides; and (b) a non-specific immune response enhancer.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion protein or polynucleotide encoding a fusion protein in combination with a physiologically acceptable carrier are provided.

Vaccines are further provided, within other aspects, comprising a fusion protein or polynucleotide encoding a fusion protein in combination with a non-specific immune response enhancer.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition or vaccine as recited above.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a breast tumor antigen, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (i) polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and (ii) complements of the foregoing polynucleotides, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the antigen from the sample.

Within related aspects, methods are provided for inhibiting the development of breast cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a breast tumor antigen, comprising contacting T cells with one or more of: (i) a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions,

deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and complements of the foregoing polynucleotides; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of breast cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of breast cancer in a patient, comprising the steps of: (a) incubating $CD4^+$ and/or CD8+ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of an breast tumor antigen or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and complements of such polynucleotides; (ii) a polynucleotide encoding such a polypeptide; or (iii) an antigen-presenting cell that expresses such a polypeptide; such that T cells proliferate; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of breast cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present provides methods for determining the presence or absence of a cancer in a patient, comprising (a) contacting a biological

sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be breast cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor antigen, wherein the antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (i) polynucleotides recited in Figure 1 or Figure 2; and (ii) complements of the foregoing polynucleotides; (b) detecting in the sample a level of a polynucleotide that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a

polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor antigen, wherein the antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (i) polynucleotides recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186; and (ii) complements of the foregoing polynucleotides; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1L depict partial sequences of representative polynucleotides encoding breast tumor antigens (SEQ ID NOs: 1-35), and the predicted amino acid sequences of the encoded polypeptides (SEQ ID NOs: 58-77).

Figures 2A-2I depict partial sequences of representative polynucleotides encoding further breast tumor antigens (SEQ ID NOs: 36-57), and the predicted amino acid sequences of the encoded polypeptides (SEQ ID NOs: 78-89).

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for the therapy and diagnosis of cancer, such as breast cancer. The compositions described herein may include immunogenic polypeptides, nucleic acid sequences encoding such polypeptides, binding agents such as antibodies that bind to a polypeptide, antigen presenting cells (APCs) that express such a polypeptide and/or immune system cells (*e.g.*, T cells). Polypeptides of the present invention generally comprise at least an immunogenic portion of a breast tumor antigen or a variant thereof. A "breast tumor antigen" is a protein that is expressed by breast tumor cells (preferably human cells) and that reacts detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with breast cancer. Nucleic acid sequences of the subject invention generally comprise a DNA or RNA sequence that encodes all or a portion of such a polypeptide, or that is complementary to such a sequence. Antibodies are generally immune system proteins, or antigen-binding fragments thereof, that are capable of binding to a portion of a polypeptide as described above. Antigen presenting cells include dendritic cells and macrophages that express a polypeptide as described above. T cells that may be employed within such compositions are generally T cells that are specific for a polypeptide as described above.

The present invention is based on the discovery of previously unknown human breast tumor antigens. Partial sequences of polynucleotides encoding specific breast tumor antigens are provided in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186. Certain extended sequences are also provided in SEQ ID NOs:187-192.

BREAST TUMOR ANTIGEN POLYNUCLEOTIDES

Any polynucleotide that encodes a breast tumor antigen or a portion or other variant thereof as described herein is encompassed by the present invention. Preferred polynucleotides comprise at least 15 consecutive nucleotides, and preferably at least 30 consecutive nucleotides, that encode a portion of a breast tumor antigen. More preferably, a polynucleotide encodes an immunogenic portion of a breast tumor

antigen. Polynucleotides complementary to any such sequences are also encompassed by the present invention. Polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a breast tumor antigen or a portion thereof) or may comprise a variant of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native breast tumor antigen. The effect on the immunogenicity of the encoded polypeptide may generally be assessed as described herein. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a polynucleotide sequence that encodes a native breast tumor antigen or a portion thereof. The percent identity may be readily determined by comparing sequences using computer algorithms well known to those of ordinary skill in the art, such as Megalign, using default parameters. Certain variants are substantially homologous to a native gene, or a portion or complement thereof. Such polynucleotide variants are capable of hybridizing under moderately stringent conditions to a naturally occurring DNA sequence encoding a native breast tumor antigen (or a complementary sequence). Suitable moderately stringent conditions include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides

that vary due to differences in codon usage are specifically contemplated by the present invention.

Polynucleotides may be prepared using any of a variety of techniques. For example, a polynucleotide may be identified, as described in more detail below, by screening a breast tumor cDNA expression library with sera of patients with breast cancer. Alternatively, polypeptides may be amplified from cDNA prepared from breast tumor cells. Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this approach, sequence-specific primers may be designed based on the sequences provided herein, and may be purchased or synthesized.

An amplified portion may be used to isolate a full length gene from a suitable library (*e.g.,* a breast tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.,* by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see* Sambrook et al., *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences are then assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures of about 68°C to 72°C. The amplified region may be sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (*see* Triglia et al., *Nucl. Acids Res. 16*:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res. 19*:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be

performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence.

Certain nucleic acid sequences of cDNA molecules encoding portions of breast tumor antigens are provided in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186. These polynucleotides were isolated by serological screening of a breast tumor cDNA expression library. The library was prepared from unamplified cDNA derived from a pool of three human breast tumors grown in a SCID mouse, in the vector pScreen. Sera from two human patients with breast cancer were pooled for the screen. The polynucleotides recited herein, as well as full length polynucleotides comprising such sequences, other portions of such full length polynucleotides, and sequences complementary to all or a portion of such full length molecules, are specifically encompassed by the present invention.

Polynucleotide variants may generally be prepared by any method known in the art, including chemical synthesis by, for example, solid phase phosphoramidite chemical synthesis. Modifications in a polynucleotide sequence may also be introduced using standard mutagenesis techniques, such as oligonucleotide-directed site-specific mutagenesis (*see* Adelman et al., *DNA* 2:183, 1983). Alternatively, RNA molecules may be generated by *in vitro* or *in vivo* transcription of DNA sequences encoding a breast tumor antigen, or portion thereof, provided that the DNA is incorporated into a vector with a suitable RNA polymerase promoter (such as T7 or SP6). Certain portions may be used to prepare an encoded polypeptide, as described herein. In addition, or alternatively, a portion may be administered to a patient such that the encoded polypeptide is generated *in vivo* (*e.g.*, by transfecting antigen-presenting cells such as dendritic cells with a cDNA construct encoding a breast tumor antigen polypeptide, and administering the transfected cells to the patient).

A portion of a sequence complementary to a coding sequence (*i.e.*, an antisense polynucleotide) may also be used as a probe or to modulate gene expression. cDNA constructs that can be transcribed into antisense RNA may also be introduced into cells of tissues to facilitate the production of antisense RNA. An antisense polynucleotide may be used, as described herein, to inhibit expression of a breast tumor

antigen. Antisense technology can be used to control gene expression through triple-helix formation, which compromises the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules (*see* Gee et al., *In* Huber and Carr, *Molecular and Immunologic Approaches*, Futura Publishing Co. (Mt. Kisco, NY; 1994)). Alternatively, an antisense molecule may be designed to hybridize with a control region of a gene (*e.g.*, promoter, enhancer or transcription initiation site), and block transcription of the gene; or to block translation by inhibiting binding of a transcript to ribosomes.

A portion of a coding sequence or a complementary sequence may also be designed as a probe or primer to detect gene expression. Probes may be labeled by a variety of reporter groups, such as radionuclides and enzymes, and are preferably at least 10 nucleotides in length, more preferably at least 20 nucleotides in length and still more preferably at least 30 nucleotides in length. Primers, as noted above, are preferably 22-30 nucleotides in length.

Any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl- methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

Nucleotide sequences as described herein may be joined to a variety of other nucleotide sequences using established recombinant DNA techniques. For example, a polynucleotide may be cloned into any of a variety of cloning vectors, including plasmids, phagemids, lambda phage derivatives and cosmids. Vectors of particular interest include expression vectors, replication vectors, probe generation vectors and sequencing vectors. In general, a vector will contain an origin of replication functional in at least one organism, convenient restriction endonuclease sites and one or more selectable markers. Other elements will depend upon the desired use, and will be apparent to those of ordinary skill in the art.

Within certain embodiments, polynucleotides may be formulated so as to permit entry into a cell of a mammal, and expression therein. Such formulations are particularly useful for therapeutic purposes, as described below. Those of ordinary skill in the art will appreciate that there are many ways to achieve expression of a polynucleotide in a target cell, and any suitable method may be employed. For example, a polynucleotide may be incorporated into a viral vector such as, but not limited to, adenovirus, adeno-associated virus, retrovirus, or vaccinia or other pox virus (*e.g.*, avian pox virus). Techniques for incorporating DNA into such vectors are well known to those of ordinary skill in the art. A retroviral vector may additionally transfer or incorporate a gene for a selectable marker (to aid in the identification or selection of transduced cells) and/or a targeting moiety, such as a gene that encodes a ligand for a receptor on a specific target cell, to render the vector target specific. Targeting may also be accomplished using an antibody, by methods known to those of ordinary skill in the art.

Other formulations for therapeutic purposes include colloidal dispersion systems, such as macromolecule complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, and liposomes. A preferred colloidal system for use as a delivery vehicle *in vitro* and *in vivo* is a liposome (*i.e.*, an artificial membrane vesicle). The preparation and use of such systems is well known in the art.

BREAST TUMOR ANTIGEN POLYPEPTIDES

Within the context of the present invention, polypeptides may comprise at least an immunogenic portion of a breast tumor antigen or a variant thereof, as described herein. As noted above, a "breast tumor antigen" is a protein that is expressed by breast tumor cells and that reacts detectably within an immunoassay (such as an ELISA) with antisera from a patient with breast cancer. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

An "immunogenic portion," as used herein is a portion of an antigen that is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of a breast tumor antigen or a variant thereof. Preferred immunogenic portions are encoded by sequences recited in Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) and SEQ ID NOs:90-186, or the complements thereof. Further immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the antigen in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native breast carcinoma antigen is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the reactivity of the full length polypeptide. Such screens may generally be performed using methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ^{125}I-labeled Protein A.

As noted above, a composition may comprise a variant of a native breast tumor antigen. A polypeptide "variant," as used herein, is a polypeptide that differs from a native breast tumor antigen in one or more substitutions, deletions, additions and/or insertions, such that the immunogenicity of the polypeptide is not substantially

diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native antigen, or may be diminished by less than 50%, and preferably less than 20%, relative to the native antigen. Such variants may generally be identified by modifying one of the above polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein.

Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (*e.g.*, poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast and higher eukaryotic cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having fewer than about 100 amino acids, and generally fewer than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See* Merrifield, *J. Am. Chem. Soc. 85*:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Applied BioSystems, Inc. (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises one polypeptide as described herein and a known tumor antigen, such as a breast tumor antigen, or a variant of such an antigen. Fusion proteins may generally be prepared using standard techniques. For example, a fusion protein may be prepared recombinantly. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of

the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and the second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene 40*:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA 83*:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

Fusion proteins are also provided that comprise a polypeptide of the present invention together with an unrelated immunogenic protein. Preferably the

immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute et al. *New Engl. J. Med.*, *336*:86-91, 1997).

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

BINDING AGENTS

The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to a breast tumor antigen. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a breast tumor antigen if it reacts at a detectable level (within, for example, an ELISA) with a breast tumor antigen, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two separate molecules such that a "complex" is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex formation exceeds about 10^3 L/mol. The binding constant maybe determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as breast cancer, using the representative assays provided herein. In other words, antibodies or other binding agents that bind to a breast tumor antigen will generate a signal indicating the presence of a cancer in at least about

20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, urine and/or tumor biopsies) from patients with and without breast cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, and RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.,* Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for the antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol. 6*:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested

by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y, ^{123}I, ^{125}I, ^{131}I, ^{186}Re, ^{188}Re, ^{211}At, and ^{212}Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diptheria toxin, cholera toxin, gelonin, Pseudomonas exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers which provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T cells specific for a breast tumor protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be present within (or isolated from) bone marrow, peripheral blood or a fraction of bone marrow or peripheral blood of a mammal, such as a patient, using a commercially available cell separation system, such as the CEPRATE™ system, available from CellPro Inc., Bothell WA (see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human animals, cell lines or cultures.

T cells may be stimulated with a breast tumor polypeptide, polynucleotide encoding a breast tumor polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide. Preferably, a breast tumor polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a breast tumor antigen polypeptide if the T cells kill target cells coated with a breast tumor antigen polypeptide or expressing a gene encoding such a polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen et al., *Cancer Res.*

54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a breast tumor antigen polypeptide (200 ng/ml - 100 μg/ml, preferably 100 ng/ml - 25 μg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells and/or contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan et al., Current Protocols in Immunology, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a breast tumor antigen polypeptide, polynucleotide or polypeptide-expressing APC may be $CD4^+$ and/or $CD8^+$. Breast tumor antigen-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient or a related or unrelated donor and are administered to the patient following stimulation and expansion.

For therapeutic purposes, $CD4^+$ or $CD8^+$ T cells that proliferate in response to a breast tumor antigen polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a breast tumor antigen, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a breast tumor antigen polypeptide. Alternatively, one or more T cells that proliferate in the presence of a breast tumor antigen can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

PHARMACEUTICAL COMPOSITIONS AND VACCINES

Within certain aspects, polypeptides, polynucleotides and/or binding agents may be incorporated into pharmaceutical compositions or vaccines. Pharmaceutical compositions comprise one or more such compounds and a

physiologically acceptable carrier. Vaccines may comprise one or more such compounds and a non-specific immune response enhancer. A non-specific immune response enhancer may be any substance that enhances an immune response to an exogenous antigen. Examples of non-specific immune response enhancers include adjuvants, biodegradable microspheres (*e.g.,* polylactic galactide) and liposomes (into which the compound is incorporated; *see e.g.,* Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and vaccines within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion polypeptide or as a separate compound within the composition or vaccine.

A pharmaceutical composition or vaccine may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems 15*:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface. In a preferred embodiment, the DNA may be introduced using a viral expression system (*e.g.,* vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA 86*:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci. 569*:86-103, 1989; Flexner et al., *Vaccine 8*:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973;

U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques 6*:616-627, 1988; Rosenfeld et al., *Science 252*:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA 91*:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA 90*:11498-11502, 1993; Guzman et al., *Circulation 88*:2838-2848, 1993; and Guzman et al., *Cir. Res. 73*:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer et al., *Science 259*:1745-1749, 1993 and reviewed by Cohen, *Science 259*:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (*e.g.,* polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268 and 5,075,109.

Such compositions may also comprise buffers (*e.g.,* neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.,* glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, chelating agents such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide) and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of non-specific immune response enhancers may be employed in the vaccines of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bortadella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI), Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ), alum, biodegradable microspheres, monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the vaccines provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.,* IFN-γ, IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.,* IL-4, IL-5, IL-6, IL-10 and TNF-β) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol. 7*:145-173, 1989.

Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Ribi ImmunoChem Research Inc. (Hamilton, MT; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). Also preferred is AS-2 (SmithKline Beecham). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555.

Another preferred adjuvant is a saponin, preferably QS21, which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprises an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210. Any vaccine provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a suitable carrier or excipient.

The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule or sponge that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology and administered by, for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane. Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and vaccines to facilitate production of an immune response. Delivery vehicles include antigen presenting cells, such as dendritic cells and macrophages. Such cells may be transfected with a polynucleotide encoding a breast tumor antigen (or portion or other variant thereof) such that the breast tumor antigen polypeptide is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for

therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. I*n vivo* and *ex vivo* transfection of dendritic cells may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997.

CANCER THERAPY

In further aspects of the present invention, the compositions described herein may be used for immunotherapy of cancer, such as breast cancer. Within such methods, pharmaceutical compositions and vaccines are typically administered to a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and vaccines may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. Within certain preferred embodiments, a patient is afflicted with breast cancer. Such cancer may be diagnosed using criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immuno response-modifying agents (such as tumor vaccines, bacterial adjuvants and/or cytokines).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T lymphocytes (such as $CD8^+$

cytotoxic T lymphocytes and $CD4^+$ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow in vivo and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see*, for example, Cheever et al., *Immunological Reviews 157*:177, 1997).

The polypeptides provided herein may also be used to generate and/or isolate tumor-reactive T cells, which can then be administered to a patient. In one such technique, antigen-specific T cell lines may be generated by *in vivo* immunization with short peptides corresponding to immunogenic portions of the disclosed polypeptides.

The resulting antigen-specific $CD8^+$ CTL clones may be isolated from the patient, expanded using standard tissue culture techniques and returned to the patient.

Polypeptides may also be used for *ex vivo* treatment of a cancer, such as breast cancer. For example, cells of the immune system, such as T cells, may be isolated from the peripheral blood of a patient, using a commercially available cell separation system, such as CellPro Incorporated's (Bothell, WA) CEPRATE™ system (*see* U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). The separated cells are stimulated with one or more of the immunoreactive polypeptides contained within a delivery vehicle, such as a microsphere, to provide antigen-specific T cells. The population of tumor antigen-specific T cells is them expanded using standard techniques and the cells may be administered back to the patient as described, for example, by Chang et al., *Crit. Rev. Oncol. Hematol. 22*:213, 1996.

Within another embodiment, syngeneic or autologous dendritic cells may be pulsed with peptides corresponding to at least an immunogenic portion of a polypeptide disclosed herein. The resulting antigen-specific dendritic cells may either be transferred into a patient or employed to stimulate T cells to provide antigen-specific T cells which may, in turn, be administered to a patient. The use of peptide-pulsed dendritic cells to generate antigen-specific T cells and the subsequent use of such antigen-specific T cells to eradicate tumors in a murine model has been demonstrated by Cheever et al., *Immunological Reviews 157*:177, 1997.

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally

(*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level.. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 100 μg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a breast tumor antigen generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

METHODS FOR DETECTING CANCER

In general, a cancer may be detected in a patient based on the presence of one or more breast tumor antigens and/or polynucleotides encoding such antigens in a biological sample obtained from the patient. In other words, such antigens may be used as markers to indicate the presence or absence of a cancer such as breast cancer. In

addition, such antigens may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding an antigen, which is also indicative of the presence or absence of a cancer.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See*, *e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length breast tumor antigens and portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the antigen may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a

plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the antigen and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μg, and preferably about 100 ng to about 1 μg, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (*see*, *e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a

different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with breast cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent

groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as breast cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of breast cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without breast cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for breast cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for breast cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution

containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of breast cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1μg, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the antigens or binding agents of the present invention. The above descriptions are intended to be exemplary only.

In another embodiment, the above polypeptides may be used as markers for the progression of cancer, such as breast cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide detected by the binding agent increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide either remains constant or decreases with time.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a breast tumor antigen in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a breast tumor antigen cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the breast tumor antigen. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a breast tumor antigen may be used in a hybridization assay to detect the presence of polynucleotide encoding the antigen in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a breast tumor antigen that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in any one of Figures 1A-1L (SEQ ID NOs: 1-35), Figures 2A-2I (SEQ ID NOs: 36-57) or SEQ ID NOs:90-186. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, *51*:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a sample tissue and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on

samples obtained from biological samples taken from a test patient and an individual who is not afflicted with breast cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple breast tumor antigen markers may be assayed within a given sample. It will be apparent that binding agents specific for different antigens provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of antigen markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for antigens provided herein may be combined with assays for other known tumor antigens.

DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a breast tumor antigen. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a breast tumor antigen in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a breast tumor antigen. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a breast tumor antigen.

The following Example is offered by way of illustration and not by way of limitation.

EXAMPLE

Identification of Breast Tumor Antigen cDNAs

This Example illustrates the identification of cDNA molecules encoding breast tumor antigens.

Patient sera (from two human patients with breast cancer) was adsorbed against E. *coli* and used at a 1:100 dilution in a serological expression screen performed as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989. The library screened was made from three pooled SCID-derived human breast tumors using a directional RH oligo(dT) priming cDNA library construction kit and the λScreen vector (Novagen). Approximately 600,000 pfu of the amplified library were screened, and more than 200 plaques were picked. Of 95 sequenced clones, 61 corresponded to known genes (shown in Figures 1A-1L) and 31 corresponded to novel genes (shown in Figures 2A-2I). Human nucleophosmin clones accounted for 11 of the known sequences.

Table I - Summary of Breast Tumor Antigen Partial Sequences

Sequence	Homology
BR1-1	Human calpastatin (SEQ ID NO:1; predicted amino acid sequence SEQ ID NO:58)
BR1-2	Human Nucleophosmin (Numatrin: Nucleolar phosphoprotein B23) - 11 clones (SEQ ID NO:2; predicted amino acid sequence SEQ ID NO:59)
BR1-12	Human heterogeneous ribonucleoprotein AO (SEQ ID NO:3; predicted amino acid sequence SEQ ID NO:60)
BR1-17	Human Isocitrate dehydrogenase (SEQ ID NO:4; predicted amino acid sequence SEQ ID NO:61)
BR1-18	Human RAS-related protein RAB-5A - 2 clones (SEQ ID NO:5; predicted amino acid sequence SEQ ID NO:62)
BR1-22	Human carcinoma derived Alu - ma1 - IgA FC receptor precursor (SEQ ID NO:6; predicted amino acid sequence SEQ ID NO:63)
BR1-38	Human transcription factor E2F-1 (Retinoblastoma binding protein) (SEQ ID NO:7; predicted amino acid sequence SEQ ID NO:64)

BR1-41	Human acetylglucosaminyltransferase (SEQ ID NO:8; predicted amino acid sequence SEQ ID NO:65)
BR1-43	Human ubiquitin-conjugating enzyme E2-24 kDa - 2 clones (SEQ ID NO:9; predicted amino acid sequence SEQ ID NO:66)
BR1-46	Human 28.3 kDa protein C21ORF2 (SEQ ID NO:10; predicted amino acid sequence SEQ ID NO:67)
BR1-50	Human glycogenin (primer for glycogen synthesis) (SEQ ID NO:11; predicted amino acid sequence SEQ ID NO:68)
BR1-51	Human mitosin / kinetochore protein CENP-F - 2 clones (SEQ ID NO:12; predicted amino acid sequence SEQ ID NO:69)
BR1-58	Human EIF-2-beta (SEQ ID NO:13; predicted amino acid sequence SEQ ID NO:70)
BR1-189	Human utrophin (SEQ ID NO:23; predicted amino acid sequence SEQ ID NO:73)
BR1-194	HMG-17, non histone chromosomal protein (SEQ ID NO:24; predicted amino acid sequence SEQ ID NO:74)
BR1-201	Human xs99 mRNA (SEQ ID NO:25; predicted amino acid sequence SEQ ID NO:75)
BR1-210	Human beta filamin (SEQ ID NO:26)
BR1-213	Human farnesyl pyrophosphate synthetase (SEQ ID NO:27)
BR1-219	Murine type C retrovirus (SEQ ID NO:28)
BR1-220	Human proto-oncogene (SEQ ID NO:29)
BR1-221	Human Rad 50 (SEQ ID NO:30)
BR1-222	Human placental protein 15 (SEQ ID NO:31)
BR1-92	Human microfibrillar associated protein (SEQ ID NO:14)
BR1-97	Human torsin A (3' UTR) (SEQ ID NO:15)
BR1-98	Human MLN 50; overexpressed in breast carcinoma (3' UTR) (SEQ ID NO:16)
BR1-107	Human ERK activator kinase (SEQ ID NO:17)
BR1-110	Human Alanyl-tRNA synthase (SEQ ID NO:18)
BR1-113	Human arginine methyl transferase (SEQ ID NO:19)
BR1-119	Human artemin neurotrophic factor (SEQ ID NO:20; predicted amino acid sequence SEQ ID NO:71)
BR1-123	Human ubiquitin carrier protein E2EPF (SEQ ID NO:21)
BR1-224	Human NADH ubiquinone oxyreductase chain 4 (SEQ ID NO:32)

BR1-232	Human HIV 1 inducer of short transcripts binding protein (TTF-1 interacting peptide 21; SEQ ID NO:33; predicted amino acid sequences SEQ ID NOs: 76 and 77)
BR1-234	Human receptor-like furin (3'UTR); (SEQ ID NO:34)
BR1-237	Human riboprotein S12 (SEQ ID NO:35)
BR1-124	Human cDNA YH95A05; similar to human and mouse TBX2 protein (SEQ ID NO:22; predicted amino acid sequence SEQ ID NO:72)
BR1-109	Human unknown KIAA0092 (SEQ ID NO:45; predicted amino acid sequence SEQ ID NO:84)
BR1-16	Novel (SEQ ID NO:36)
BR1-45	Novel; homology to ESTs - 2 clones (SEQ ID NO:37; predicted amino acid sequence SEQ ID NO:78)
BR1-48	Novel; homology to ESTs (SEQ ID NO:38; predicted amino acid sequence SEQ ID NO:79)
BR1-49	Novel; homology to ESTs (SEQ ID NO:39)
BR1-52	Novel; homology to ESTs, human signal recognition particle receptor beta subunit (SEQ ID NO:40; predicted amino acid sequence SEQ ID NO:80)
BR1-72	Novel; homology to ESTs (SEQ ID NO:49)
BR1-74	Novel; homology to mouse NG27, angiopoietin-2 and ESTs (SEQ ID NO:48; predicted amino acid sequence SEQ ID NO:87)
BR1-77	Novel (SEQ ID NO:50)
BR1-82	Novel; homology to ESTs (SEQ ID NO:51)
BR1-188	Novel; homology to ESTs (SEQ ID NO:47; predicted amino acid sequence SEQ ID NO:86)
BR1-204	Novel; homology to ESTs (SEQ ID NO:54; predicted amino acid sequence SEQ ID NO:89)
BR1-207	Novel; homology to ESTs (SEQ ID NO:55)
BR1-214	Novel (SEQ ID NO:56)
BR1-215	Novel (SEQ ID NO:57)
BR1-85	Novel; homology to human lipoprotein receptor related protein 5 and LDL receptor member LR3, and to EST (SEQ ID NO:52)
BR1-90	Novel; homology to ESTs and TGF-beta binding protein (SEQ ID NO:53; predicted amino acid sequence SEQ ID NO:88)
BR1-91	Novel; homology to ESTs and *C. elegans* hypothetical (SEQ ID NO:41; predicted amino acid sequence SEQ ID NO:81)

BR1-95	Novel; homology to ESTs - 4 clones (SEQ ID NO:42)
BR1-102	Novel; homology to ESTs and *S. pombe* hypothetical (SEQ ID NO:43; predicted amino acid sequence SEQ ID NO:82)
BR1-105	Novel; homology to ESTs and tumor suppressor MN1 (SEQ ID NO:44; predicted amino acid sequence SEQ ID NO:83)
BR1-111	Novel (SEQ ID NO:46; predicted amino acid sequence SEQ ID NO:85); similar to chicken inner centromere protein

Further sequenced clones are summarized in Table II.

[NOTE - THE SEQUENCE LISTING YOU PROVIDED APPEARS TO HAVE THE BR2-23 AND BR2-24 SEQUENCES REPEATED IN PLACE OF SEQUENCES FOR BR2-79 AND BR2-80. WE DO NOT APPEAR TO HAVE SEQUENCES FOR BR2-79 OR BR2-80. PLEASE PROVIDE THESE SEQUENCES.]

Table II - Summary of Breast Tumor Antigen Partial Sequences

Sequence	SEQ ID NO	Homology
BR1-133	90	human clone 327 J 16 and ESTs
BR1-134	91	human clone PAC 121 G13 and ESTs
BR1-135	92	novel
BR1-136	93	human clone hrpk.401-G-18
BR1-137	94	novel; homology to ESTs
BR1-143	95	human CAGH32 mRNA
BR1-147	96	novel; homology to EST
BR1-149	97	human anti oncogen and ESTs
BR1-151	98	novel; homology to ESTs
BR1-152	99	human TACC1 mRNA and ESTs
BR1-153	100	human helix-loop-helix proteins and ESTs
BR1-155	101	probable tumor suppressor protein; human NUP98-NUP96 precursor and ESTs
BR1-158	102	human DNA sequence from clone 118

BR1-160	103	human mitochondrial genome; cytochrome C oxidase polypeptide and ESTs
BR1-163	104	novel; homology to ESTs
BR1-164	105	T-complex protein and ESTs
BR1-166	106	clone hRP 161 and ESTs
BR1-172	107	RHOA multi-drug resistance protein and ESTs
BR2-2	108	80K-L protein and ESTs
BR2-4	109	caseine kinase II alpha subunit and ESTs
BR2-7	110	novel; homology to ESTs
BR2-8	111	novel; homology to ESTs
BR2-9	112	Podocalyxin-like (PODXL); EST
BR2-10	113	human hnRNP core protein A; ribonucleoprotein A1; ESTs
BR2-13	114	mitochondrial DNA; ATPase subunit 6; ESTs
BR2-16	115	novel; homology to ESTs
BR2-17	116	novel
BR2-18	117	novel; homology to ESTs
BR2-19	118	Yamaguchi viral oncogene; EST
BR2-23	119	polypyrimidine binding protein; ESTs
BR2-24	120	novel; ESTs
BR2-25	121	sex determining region Y; transcription factor SOX-9
BR2-26	122	AND(P) H:oxireductase gene; ESTs
BR2-27	123	human GSA mRNA; GTP-binding regulatory protein; ESTs
BR2-29	124	unknown RG459N13; ESTs
BR2-30	125	human DNA clone 1191B2; ESTs
BR2-31	126	novel; homology to ESTs
BR2-33	127	human DNA clone 850H21; EST
BR2-35	128	human CpG island DNA; ESTs
BR2-36	129	novel; homology to ESTs
BR2-39	130	novel; homology to ESTs
BR2-40	131	novel; homology to ESTs
BR2-41	132	novel; homology to EST
BR2-42	133	human mitochondrial DNA; ESTs
BR2-43	134	unknown BAC 215012; ESTs
BR2-44	135	human mitochondrial DNA; ESTs
BR2-49	136	28 kDa heat shock protein; 27 kD protein 1; ESTs

BR2-51	137	clone DJ0726N20
BR2-52	138	importin beta-2 subunit; ESTs
BR2-53	139	novel
BR2-54	140	diphosphoinositol polyphosphate; ESTs
BR2-55	141	novel; ESTs
BR2-56	142	cytochrome-c oxidase; ESTs
BR2-58	143	Phosphoglycerate dehydrogenase; ESTs
BR2-59	144	human RNA hMCM2; DNA replication licensing factor; ESTs
BR2-65	145	ribosomal protein S7 RNA; ESTs
BR2-66	146	ESTs
BR2-67	147	phospholipid peroxidase; ESTs
BR2-68	148	CGI-45 protein; ESTs
BR2-69	149	human protein kinase; ESTs
BR2-70	150	proteasome alpha 2 subunit; ESTs
BR2-71	151	deoxycytidine kinase mRNA; ESTs
BR2-72	152	novel
BR2-75	153	arginine/serine-rich 7 gene; ESTs
BR2-79	154	novel
BR2-80	155	novel; ESTs
BR2-81	156	ribosomal protein L35 mRNA; ESTs
BR2-85	157	novel; ESTs
BR2-87	158	human HSPCO39 protein mRNA; ESTs
BR2-88	159	human zinc finger protein; ESTs
BR2-89	160	novel; ESTs
BR2-90	161	novel; ESTs
BR2-91	162	90 kD heat shock protein; ESTs
BR2-92	163	human heterogeneous RNA; ESTs
BR2-98	164	ESTs
BR2-101	165	chromosome 17; ESTs
BR2-104	166	adenocarcinoma antigen; major gastrointestinal protein; ESTs
BR2-105	167	novel; homology to ESTs
BR2-106	168	novel; homology to ESTs
BR2-107	169	human KIAA1004 protein; ESTs
BR2-108	170	human alpha-CP1 mRNA; hnRNP protein x homolog; ESTs
BR2-109	171	multispanning protein; ESTs

BR2-118	172	chromosome 17; ESTs
BR2-122	173	novel; homology to ESTs
BR2-125	174	novel; homology to ESTs
BR2-127	175	malate dehydrogenase; ESTs
BR2-128	176	novel
BR2-131	177	novel
BR2-145	178	complement C1Q subunit
BR2-146	179	novel; homology to ESTs
BR2-149	180	ribonucleoprotein U; scaffold attachment factor; ESTs
BR2-152	181	human proteasome subunit; ESTs
BR2-156	182	human casein kinase 1; ESTs
BR2-158	183	novel; homology to ESTs
BR2-159	184	human mitochondrion; cytochrome c oxidase; ESTs
BR2-163	185	novel
BR2-167	186	unknown BAC GS368F15; ESTs

Full insert sequences were also obtained for the sequences shown in Table III.

Table III - Full Length Insert Sequences for Breast Tumor Antigens

Sequence	SEQ ID NO	Homology
BR1-90	187	Novel; homology to ESTs
BR1-91	188	Novel; homology to EST
BR1-85	189	Novel; homology to ESTs
BR1-77	190	Novel; homology to ESTs
BR1-105	191	Novel; homology to ESTs
BR1-204	192	Novel; homology to ESTs

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration,

various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

1. An isolated polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) polynucleotides recited in any one of SEQ ID NOs:36-57, 90-94, 96, 98, 102, 104, 110, 111, 115-117, 120, 124-127, 129-135, 137, 139, 141, 146, 152, 154, 155, 157, 160, 161, 164, 165, 167-169, 172-174, 176, 177, 179, 183, 185 or 186; and

(b) complements of the foregoing polynucleotides.

2. A polypeptide according to claim 1, wherein the polypeptide comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) polynucleotides recited in any one of SEQ ID NOs:36-57, 90-94, 96, 98, 102, 104, 110, 111, 115-117, 120, 124-127, 129-135, 137, 139, 141, 146, 152, 154, 155, 157, 160, 161, 164, 165, 167-169, 172-174, 176, 177, 179, 183, 185 or 186; and

(b) complements of such polynucleotides.

3. An isolated polynucleotide encoding at least 15 amino acid residues of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence designated as BR1-16, BR1-77, BR1-214, BR1-215, BR1-111, BR1-135, BR1-116, BR2-17, BR2-53, BR2-72, BR2-79, BR2-128, BR2-131 or BR2-163 (SEQ ID NOs:36, 50, 56, 57, 46, 92, 116, 139, 152, 154, 176, 177 and 185) or a complement of any of the foregoing sequences.

4. A polynucleotide encoding a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the

ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NOs:36-57, 92, 94, 96, 98, 104, 110, 111, 115-117, 120, 126, 129-132, 139, 141, 146, 152, 154, 155, 157, 160, 161, 164, 167, 168, 173-174, 176, 177, 179, 183 or 185, or a complement of any of the foregoing sequences.

5. An isolated polynucleotide complementary to a polynucleotide according to claim 3 or claim 4.

6. An expression vector comprising a polynucleotide according to a claim 3 or claim 4.

7. An expression vector comprising a polynucleotide according claim 5.

8. A host cell transformed or transfected with an expression vector according to claim 7.

9. A pharmaceutical composition comprising:

(a) a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs:36-57, 90-94, 96, 98, 102, 104, 110, 111, 115-117, 120, 124-127, 129-135, 137, 139, 141, 146, 152, 154, 155, 157, 160, 161, 164, 165, 167-169, 172-174, 176, 177, 179, 183, 185 or 186; and

(ii) complements of the foregoing polynucleotides; and

(b) a physiologically acceptable carrier.

10. A vaccine comprising:

(a) a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs:36-57, 90-94, 96, 98, 102, 104, 110, 111, 115-117, 120, 124-127, 129-135, 137, 139, 141, 146, 152, 154, 155, 157, 160, 161, 164, 165, 167-169, 172-174, 176, 177, 179, 183, 185 or 186; and

(ii) complements of the foregoing polynucleotides; and

(b) a non-specific immune response enhancer.

11. A pharmaceutical composition comprising:

(a) a polynucleotide encoding at least 15 amino acid residues of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs:36-57, 90-94, 96, 98, 102, 104, 110, 111, 115-117, 120, 124-127, 129-135, 137, 139, 141, 146, 152, 154, 155, 157, 160, 161, 164, 165, 167-169, 172-174, 176, 177, 179, 183, 185 or 186; and

(ii) complements of the foregoing polynucleotides; and

(b) a physiologically acceptable carrier.

12. A vaccine comprising:

(a) a polynucleotide encoding at least 15 amino acid residues of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino

acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs:36-57, 90-94, 96, 98, 102, 104, 110, 111, 115-117, 120, 124-127, 129-135, 137, 139, 141, 146, 152, 154, 155, 157, 160, 161, 164, 165, 167-169, 172-174, 176, 177, 179, 183, 185 or 186; and

(ii) complements of the foregoing polynucleotides; and

(b) a non-specific immune response enhancer.

13. A pharmaceutical composition comprising:

(a) an antibody or antigen-binding fragment thereof that specifically binds to a breast tumor antigen, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs:36-57, 90-94, 96, 98, 102, 104, 110, 111, 115-117, 120, 124-127, 129-135, 137, 139, 141, 146, 152, 154, 155, 157, 160, 161, 164, 165, 167-169, 172-174, 176, 177, 179, 183, 185 or 186 and

(ii) complements of the foregoing polynucleotides; and

(b) a physiologically acceptable carrier.

14. A pharmaceutical composition, comprising:

(a) an antigen presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs:1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides; and

(b) a pharmaceutically acceptable carrier or excipient.

15. A pharmaceutical composition according to claim 14, wherein the antigen presenting cell is a dendritic cell or a macrophage.

16. A vaccine, comprising:

(a) an antigen presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides; and

(b) a non-specific immune response enhancer.

17. A vaccine according to claim 16, wherein the antigen presenting cell is a dendritic cell or a macrophage.

18. A polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides;

for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

19. A polynucleotide encoding at least 15 amino acid residues of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides;

for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

20. An antibody or antigen-binding fragment thereof that specifically binds to a breast tumor antigen, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides;

for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

21. An antigen presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides;

for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

22. A cell according to claim 21, wherein the antigen presenting cell is a dendritic cell or a macrophage.

23. A cell according to claim 21, wherein the antigen presenting cell is present within a vaccine according to claim 16.

24. A fusion protein comprising at least one polypeptide according to claim 1.

25. A polynucleotide encoding a fusion protein according to claim 24.

26. A pharmaceutical composition comprising a fusion protein according to claim 24 in combination with a physiologically acceptable carrier.

27. A vaccine comprising a fusion protein according to claim 24 in combination with a non-specific immune response enhancer.

28. A pharmaceutical composition comprising a polynucleotide according to claim 25 in combination with a physiologically acceptable carrier.

29. A vaccine comprising a polynucleotide according to claim 25 in combination with a non-specific immune response enhancer.

30. A pharmaceutical composition according to claim 26 or claim 28, for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

31. An effective amount of a vaccine according to claim 27 or claim 29, for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

32. A method for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a breast tumor antigen, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides;

wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the antigen from the sample.

33. A method according to claim 32, wherein the biological sample is blood or a fraction thereof.

34. A biological sample treated according to the method of claim 32, for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

35. A method for stimulating and/or expanding T cells specific for a breast tumor antigen, comprising contacting T cells with one or more of:

(i) a polypeptide comprising at least an immunogenic portion of a breast tumor antigen, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

complements of the foregoing polynucleotides;

(ii) a polynucleotide encoding such a polypeptide; and/or

(iii) an antigen presenting cell that expresses such a polypeptide;

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

36. An isolated T cell population, comprising T cells prepared according to the method of claim 35.

37. A T cell population according to claim 36, for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

38. $CD4^+$ and/or CD8+ T cells isolated from a patient and incubated with one or more of:

(i) a polypeptide comprising at least an immunogenic portion of an breast tumor antigen or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

complements of such polynucleotides;

(ii) a polynucleotide encoding such a polypeptide; or

(iii) an antigen-presenting cell that expresses such a polypeptide;

such that T cells proliferate; for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

39. A method for inhibiting the development of breast cancer in a patient, comprising the steps of:

(a) incubating $CD4^+$ and/or CD8+ T cells isolated from a patient with one or more of:

(i) a polypeptide comprising at least an immunogenic portion of an breast tumor antigen or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

complements of such polynucleotides;

(ii) a polynucleotide encoding such a polypeptide; or

(iii) an antigen-presenting cell that expresses such a polypeptide;

such that T cells proliferate;

(b) cloning one or more proliferated cells ; and

(c) administering to the patient an effective amount of the cloned T cells, and thereby inhibiting the development of breast cancer in the patient.

40. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with a binding agent that binds to a breast tumor antigen, wherein the antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(c) comparing the amount of polypeptide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

41. A method according to claim 40, wherein the binding agent is an antibody.

42. A method according to claim 41, wherein the antibody is a monoclonal antibody.

43. A method according to claim 40, wherein the cancer is breast cancer.

44. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a breast tumor antigen, wherein the antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii)complements of the foregoing polynucleotides;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polypeptide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

45. A method according to claim 44, wherein the binding agent is an antibody.

46. A method according to claim 45, wherein the antibody is a monoclonal antibody.

47. A method according to claim 44, wherein the cancer is breast cancer.

48. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor antigen, wherein the antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(c) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

49. A method according to claim 48, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

50. A method according to claim 48, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

51. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor antigen, wherein the antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polynucleotide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

52. A method according to claim 51, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

53. A method according to claim 51, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

54. An isolated antibody that specifically binds to a breast tumor antigen, wherein the antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides.

55. An antibody according to claim 54, wherein the antibody is a monoclonal antibody.

56. A diagnostic kit, comprising:

(a) one or more antibodies according to claim 54; and

(b) a detection reagent comprising a reporter group.

57. A kit according to claim 56, wherein the antibodies are immobilized on a solid support.

58. A kit according to claim 57, wherein the solid support comprises nitrocellulose, latex or a plastic material.

59. A kit according to claim 56, wherein the detection reagent comprises an anti-immunoglobulin, protein G, protein A or lectin.

60. A kit according to claim 56, wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.

61. An oligonucleotide comprising 10 to 40 nucleotides that hybridize under moderately stringent conditions to a polynucleotide that encodes a breast tumor antigen, wherein the antigen comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186; and

(ii) complements of the foregoing polynucleotides.

62. A oligonucleotide according to claim 61, wherein the oligonucleotide comprises 10-40 nucleotides recited in any one of SEQ ID NOs: 1-57 or 90-186.

63. A diagnostic kit, comprising:

(a) an oligonucleotide according to claim 61; and

(b) a diagnostic reagent for use in a polymerase chain reaction or hybridization assay.

BR1-1

Nucleotide Sequence

GATGACACTATCCCACCTGAATACAGACATCTCCTGGATGATAATGGACAGGACAAACCAGTGAAGCCA
CCTACAAAGAAATCAGAGGATTCAAAGAAACCTGCAGATGACCAAGACCCCATTGATGCTCTCTCAGGA
GATCTGGACAGCTGTCCCTCCACTACAGAAACCTCACAGAACACAGCAAAGGATAAGTGCAAGAAGGCT
GCTTCCAGCTCCAAAGCACCTAAGAATGGAGGTAAAGCGAAGGATTCAGCAAAGACAACAGAGGAAACT
TCCAAGCCAAAAGATGACTAAAGAAATACAAGTTAAGGTATCTGGTATCTGCATGTAAAATCTTCAGCT
GGTGGATGGTGACTTTTGAAGAACAAAAGGCTTTGGCAACAGAAAACAATTGTTCTGGGTGATTTCTAG
AATGGTTTTTGTTGAGTCTCTGAACATCCTAAATATTGGTTTGTTATTCTTTTCCAGAAAGAAAATGAA
TTTGACTGGTTCACCTGTGTACTGAGTATTGATAAACTTTGAATTTTTTTAATTGCCTTCAATTGGGAG
AGAAAGCTTTATATTTGTAAGAAATATATTTGATAAAGTTTCTTAAAGCAACACCAAAAAAACAAAAGA
AAAGCTAAGTGAATTTTTGCACATTCTACACACAGTGCCTGTAAATCTCATTTGTATTTTCAGTTTGCC
CTTAATTTTTTTTTGTTAGTGTTTAGAAAACAATGTTTTAAACATTCTTCAGTGTTCTGATTTCTTATTA
CCCCCTTTCCTCTTGGGCTTTTGAACTGTATTTGATGTTGCTTTGGGATAATGTTTATAAGTCAAACAT
AAGATATTGTACATTGGGCACATATCTCCTCTTGGGCTGCTAATAATAAATTAATAACAGGTAACCTGG
ACAAACCAGGAAGCACCAAAAAAAGCTT

Predicted Protein Sequence

DTIPPEYRHLLDDNGQDKPVKPPTKKSEDSKKPADDQDPIDALSGDLDSCPSTTETSQNTAKDKCKKAA
SSSKAPKNGGKAKDSAKTTEETSKPKDDC

BR1-2

Nucleotide Sequence

GTTGTTCTCTGGAGCAGCGTTCTTTTATCTCCGTCCGCCTTCTCTCCTACCTAAGTGCGTGCCGCCACC
CGATGGAAGATTCGATGGACATGGACATGAGCCCCCTGAGGCCCCAGAACTATCTTTTCGGTTGTGAAC
TAAAGGCCGACAAAGATTATCACTTTAAGGTGGATAATGATGAAAATGAGCACCAGTTATCTTTAAGAA
CGGTCAGTTTAGGGGCTGGTGCAAAGGATGAGTTGCACATTGTTGAAGCAGAGGCAATGAATTACGAAG
GCAGTCCAATTAAAGTAACACTGGCAACTTTGAAAATGTCTGTACAGCCAACGGTTTCCCTTGGGGGCT
TTGAAATAACACCACCAGTGGTCTTAAGGTTGAAGTGTGGTTCAGGGCCAGTGCATATTAGTGGACAGC
ACTTAGTAGCTGTGGAGGAAGATGCAGAGTCAGAAGATGAAGAGGAGGAGGATGTGAAACTCTTAAGTA
TATCTGGAAAGCGGTCTGCCCCTGGAGGTGGTAGCAAGGTTCCACAGAAAAAAGTAAAACTTGCTGCTG
ATGAAGATGATGACGATGATGATGAAGAGGATGATGATGAAGATGATGATGATGATGATTTTGATGATG
AGGAAGCTGAAGAAAAAGCGCCAGTGAAGAAATCTATACGAGATACTCCAGCCAAAAATGCACAAAAGT
CAAATCAGAATGGAAAAGACTCAAAACCATCATCAACACCAAGATCAAAAGGACAAGAATCCTTCAAGA
AACAGGAAAAAACTCCTAAAACACCAAAAGGACCTAGTTCTGTAGAAGACATTAAAGCAAAAATGCAAG
CAAGTATAGAAAAAGGTGGTTCTCTTCCCAAAGTGGAAGCCAAATTCATCAATTATGTGAAGAATTGCT
TCCGGATGACTGACCAAGAGGCTATTCAAGATCTCTGGCAGTGGAGGAAGTCTCTTTAAGAAAATAGTT
TAAACAATTTGTTAAAAAATTTTCCGTCTTATTTCATTTCTGTAACAGTTGATATCTGGCTGTCCTTTT
TATAATGCAGAGTGAGAACTTTCCCTACCGTGTTTGATAAATGTTGTCCAGGTTCTATTGCCAAGAATG
TGTTGTCCAAAATGCCGTTTAGTTTTTAAAGATGGAACTCCACCCTTTGCTTGGTTTTAAGTATGTATG
GAATGTTATGATAGGACATAGTAGTAGCGGTGGTCAGACATGGAAATGGTGGGGAGACAAAAATATACA
TGTGAAATAAAACTCAGTATTTTAATAAAGTAAAAAAAAAAAAAAAGCTT

Fig. 1A

Predicted Protein Sequence

CSLEQRSFISVRLLSYLSACRHPMEDSMDMDMSPLRPQNYLFGCELKADKDYHFKVDNDENEHQLSLRT
VSLGAGAKDELHIVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPVVLRLKCGSGPVHISGQH
LVAVEEDAESEDEEEEDVKLLSISGKRSAPGGGSKVPQKKVKLAADEDDDDDEEDDDEDDDDDFDDE
EAEEKAPVKKSIRDTPAKNAQKSNQNGKDSKPSSTPRSKGQESFKKQEKTPKTPKGPSSVEDIKAKMQA
SIEKGGSLPKVEAKFINYVKNCFRMTDQEAIQDLWQWRKSL

BR1-12

Nucleotide Sequence

AGAAGCTCTTTGTCGGAGGCCTTAAAGGAGACGTGGCTGAGGGCGACCTGATCGAGCACTTCTCGCAGT
TTGGCACCGTGGAAAAGGCCGAGATTATTGCCGACAAGCAGTCCGGCAAGAAGCGTGGATTCGGCTTCG
TGTATTTCCAGAATCACGACGCGGCAGACAAGGCCGCGGTGGTCAAGTTCCATCCGATTCAGGGCCATC
GCGTGGAGGTGAAGAAAGCAGTCCCCAAGGAGGATATCTACTCCGGTGGGGGTGGAGGCGGCTCCCGAT
CCTCCCGGGGCGGCCGAGGCGGCCGGGGGCGCGGCGGTGGTCGAGACCAGAACGGCCTTTCCAAGGGCG
GCGGCGGCGGTTACAACAGCTACGGTGGTTACGGCGGCGGCGGAGGCGGCGGCTACAATGCCTACGGAG
GCGGCGTGGCGGTTCGTCCTACGGTGGGAGCGACTACGGTAACGGCTTCGGCGGCTTCGGCAGCTACAG
CCAGCATCAGTCCTCCTATGGGCCCATGAAGAGCGGCGGCGGCGGCGGCGGTGGAGGCAGTAGCTGGGG
CGGTCGCAGTAATAGTGGACCTTACAGAGGCGGCTATGGCGGTGGGGGTGGCTATGGAGGCAGCTCCTT
CTAAAAGAAAATTTAAAATGCCTGGGAGTGGCTATAGGGGTAGCTCTTTCCAACAGCCCAAGTGGGGTC
AACTCCTAAGCCCCACCCCCTCACACACACCGCCTTCCCTGTTTTGCCCTTGGGGGAGCCACTTCTAAG
GCTGCTTACCCTTGGGGGTGTTCCTCTATTTGCCTGCCACCTCTCTTGTCTCTCCCTCTGAAGATGGAC
TCGGCCCCACATACACATTTTTGTGTTACAGTCATTGATGGACTCTATTTTTTTATTATTACTTGGACC
TTGGTCGTTTTTATACTAGCAAAATGTCTTGTTTTAATTTGTGTTTTTTGGGGGGAGGGAGGGAGTGAA
CTTGCTGATTCTGTAGCAAAACCTGGGTGGGGGTTGGGGTGGGGGGTAGTTTACTTTGTTGTAAGGACT
TGATAACCTGGCTACAGCGTTTTCTATGAAATCTACTTGGATCCCATGCCTGAAATTTGGAAGCATATG
TACAAAAATCATTTTTACGTTTTATTTTTAATAAATCATTGTGTTTGACCGTAAAAAAAAGCTT

Predicted Protein Sequence

KLFVGGLKGDVAEGDLIEHFSQFGTVEKAEIIADKQSGKKRGFGFVYFQNHDAADKAAVVKFHPIQGHR
VEVKKAVPKEDIYSGGGGGGSRSSRGGRGGRGRGGGRDQNGLSKGGGGGYNSYGGYGGGGGGGYNAYGG
GVAVRPTVGATTVTASAASAATASISPPMGP

Fig. 1B

BR1-17

Nucleotide Sequence

CGCGCTCGGACCTCTCCCGCCCTGCTCGTTCGCTCTCCAGCTTGGGATGGCCGGCTACCTGCGGGTCGT
GCGCTCGCTCTGCAGAGCCTCAGGCTCGCGGCCGGCCTGGGCGCCGGCGGCCCTGACAGCCCCCACCTC
GCAAGAGCAGCCGCGGCGCCACTATGCCGACAAAAGGATCAAGGTGGCGAAGCCCGTGGTGGAGATGGA
TGGTGATGAGATGACCCGTATTATCTGGCAGTTCATCAAGGAGAAGCTCATCCTGCCCCACGTGGACAT
CCAGCTAAAGTATTTTGACCTCGGGCTCCCAAACCGTGACCAGACTGATGACCAGGTCACCATTGACTC
TGCACTGGCCACCCAGAAGTACAGTGTGGCTGTCAAGTGTGCCACCATCACCCCTGATGAGGCCCGTGT
GGAAGAGTTCAAGCTGAAGAAGATGTGGAAAAGTCCCAATGGAACTATCCGGAACATCCTGGGGGGGAC
TGTCTTCCGGGAGCCCATCATCTGCAAAAACATCCCACGCCTAGTCCCTGGCTGGACCAAGCCCATCAC
CATTGGCAGGCACGCCCATGGCGACCAGTACAAGGCCACAGACTTTGTGGCAGACCGGGCCGGCACTTT
CAA

Predicted Protein Sequence

MAGYLRVVRSLCRASGSRPAWAPAALTAPTSQEQPRRHYADKRIKVAKPVVEMDGDEMTRIINQFIKEK
LILPHVDIQLKYFDLGLPNRDQTDDQVTIDSALATQKYSVAVKCATITPDEARVEEFKLKKMVKSPNGT
IRNILGGTVFREPIICKNIPRLVPGWTKPITIGRHAHGDQYKATDFVADRAGTF

BR1-18

Nucleotide Sequence 5'

TGGAAGTTCATTGAAGAGTCTGAAATTAGGGACTTATTTCAAATTTGGACATGGCTAGTCGAGGCGCAA
CAAGACCCAACGGGCCAAATACTGGAAATAAAATATGCCAGTTCAAACTAGTACTTCTGGGAGAGTCCG
CTGTTGGCAAATCAAGCCTAGTGCTTCGTTTTGTGAAAGGCCAATTTCATGAATTTCAAGAGAGTACCA
TTGGGGCTGCTTTTCTAACCCAAACTGTATGTCTTGATGACACTACAGTAAAGTTTGAAATATGGGATA
CAGCTGGTCAAGAACGATACCATAGCCTAGCACCAATGTACTACAGAGGAGCACAAGCAGCCATAGTTG
TATATGATATCACAAATGAGGAGTCCTTTGCAAGAGCAAAAAATTGGGTTAAAGAACTTCAGAGGCAAG
CAAGTCCTAACATTGTAATAGCTTTATCGGGAAACAAGGCCGACCTAGCAAATAAAAGAGCAGTAGATT
TCCAGGAAGCACAGTCCTATGCAGATGACAATAGTTTATTATTCATGGAGACATCCGCTAAAACATCAA
TGAATGTAAATGAAATATTCATGGCAATAGCTAAAAAATTGCCAAAGAATGAACCACAAAATCCAGGAG
CAAATTCT

Predicted Protein Sequence

MASRGATRPNGPNTGNKICQFKLVLLGESAVGKSSLVLRFVKGQFHEFQESTIGAAFLTQTVCLDDTTV
KFEIWDTAGQERYHSLAPMYYRGAQAAIVVYDITNEESFARAKNWVKELQRQASPNIVIALSGNKADLA
NKRAVDFQEAQSYADDNSLLFMETSAKTSMNVNEIFMAIAKKLPKNEPQNPGANS

Fig. 1C

BR1-22

Nucleotide Sequence

GTCAGGGTCAGATCATATGAATCCCAAATGGTTATCCGTCCCCACAAGTCATTTGATGAGAATGGCTTT
GACTACTTACTAACATACAGTGACAATCCCCAAACGGTGTTTCCTCGCTACTGTGTTAGTTGGATGGTT
TCCAGTGGCATGCCAGATTTCCTGGAGAAGCTGCACATGGCCACTCTGAAAGCCAAGAATATGGAGATT
AAAGTAAAGGACTACATCTCAGCTAAGCCTCTGGAAATGAGTAGTGAAGCCAAGGCCACCAGCCAGTCC
TCTGAGCGAAAGAACGAGGGCAGCTGTGGCCCTGCTCGGATTGAGTATGCTTGACAGGCTTTGGGATAA
GAAGGGACAAGGTGCTTCTAGCCCTGTCTCAGTCCGTTATCACTCTGCTGTAGAAGGGGGACATGCCAC
ATGTATTAGAAGGCATCTGCTGTAACTTCCAGTGCAAGATAATTCAATAACTGATGTCCCATTTCATTC
AGAGCCCTTATTGCTCTTATCAAAACAGAAGAAGGCTACATTTGTGGGAGTGTTGTCATATTCTCAGGC
CAACTGTTTTGAAATTCGGTATCTCACTGAGCTAATCTGGAACAAACCTCTCACCTCAGGCCAGAAGGG
GAT

Predicted Protein Sequence

VRVRSYESQMVIRPHKSFDENGFDYLLTYSDNPQTVFPRYCVSWMVSSGMPDFLEKLHMATLKAKNMEI
KVKDYISAKPLEMSSEAKATSQSSERKNEGSCGPARIEYA

BR1-38

Nucleotide Sequence 5'

GGACGCCAGCGCCCCGCCGGCTCCCACCGGCCTCGCGGCGCCCGCCGCCGGCCCCTGCGACCCTGACCT
GCTGCTCTTCGCCACACCGCAGGCGCCCCGGCCCACACCCAGTGCGCCGCGGCCCGCGCTCGGCCGCCC
GCCGGTGAAGCGGAGGCTGGACCTGGAAACTGACCATCAGTACCTGGCCGAGAGCAGTGGGCCAGCTCG
GGGCAGAGGCCGCCATCCAGGAAAAGGTGTGAAATCCCCGGGGGAGAAGTCACGCTATGAGACCTCACT
GAATCTGACCACCAAGCGCTTCCTGGAGCTGCTGAGCCACTC

Predicted Protein Sequence

DASAPPAPTGLAAPAAGPCDPDLLLFATPQAPRPTPSAPRPALGRPPVKRRLDLETDHQYLAESSGPAR
GRGRHPGKGVKSPGEKSRYETSLNLTTKRFLELLSH

BR1-41

Nucleotide Sequence 3'

AGGAGACGGCCCAGGCCATCGCCTCCTACGGCAGCGCGGTCACGCACATCCGGCAGCCCGACCTGAGCA
GCATTGCGGTGCCGCCGGACCACCGCAAGTTCCAGGGCTACTACAAGATCGCGCGCCACTACCGCTGGG
CGCTGGGCCAGGTCTTCCGGCAGTTTCGCTTCCCCGCGGCCGTGGTGGTGGAGGATGACCTGGAGGTGG
CCCCGGACTTCTTCGAGTACTTTCGGGCCACCTATCCGCTGCTGAAGGCCGACCCCTCCCTGTGGTGCG
TCTCGGCCTGGAATGACAACGGCAAGGAGCAGATGGTGGACGCCAGCAGGCCTGAGCTGCTCTACCGCA
CCGACTTTTTCCCTGGCCTGGGCTGGCTGCTGTTGGCCGAGCTCTGGGCTGAGCTGGAGCCCAAGTGGC
CAAAGGCCTTCTGGGACGACTGGATGCGGCGGCCGGAGCAGCGGCAGGGGCGGGCCTGCATACGCCCTG
AGATCTCAAGAACGATGACCTTTGGCCGCAAGGGTGTGAGCCACGGGCAGTT

Fig. 1D

Predicted Protein Sequence

ETAQAIASYGSAVTHIRQPDLSSIAVPPDHRKFQGYYKIARHYRWALGQVFRQFRFPAAVVVEDDLEVA
PDFFEYFRATYPLLKADPSLWCVSAWNDNGKEQMVDASRPELLYRTDFFPGLGWLLLAELWAELEPKWP
KAFWDDWMRRPEQRQGRACIRPEISRTMTFGRKGVSHGQ

BR1-43

Nucleotide Sequence

GCGGCGGGTTGGTCTACGCTGTGCGCGGCGGACGTCGGAGGCAGCGGGGAGCGGAGCGGGGCCGCCGGG
GCCTCTCCAGGGCCGCAGCGGCAGCAGTTGGGCCCCCCGCCCCGGCCGGCGGACCGAAGAACGCAGGAA
GGGGGCCGGGGGGACCCGCCGCCGGCCGGCCGCAGCCATGAACTCCAACGTGGAGAACCTACCCCCGCA
CATCATCCGCCTGGTGTACAAGGAGGTGACGACACTGACCGCAGACCCACCCGATGGCATCAAGGTCTT
TCCCAACGAGGAGGACCTCACCGACCTCCAGGTCACCATCGAGGGCCCTGAGGGGACCCCATATGCTGG
AGGTCTGTTCCGCATGAAACTCCTGCTGGGGAAGGACTTCCCTGCCTCCCCACCCAAGGGCTACTTCCT
GACCAAGATCTTCCACCCGAACGTGGGCGCCAATGGCGAGATCTGCGTCAACGTGCTCAAGAGGGACTG
GACGGCTGAGCTGGGCATCCGACACGTACTGCTGACCATCAAGTGCCTGCTGATCCACCCTAACCCCGA
GTCTGCACTCAACGAGGAGGCGGGCCGCCTGCTCTTGGAGAACTACGAGGAGTATGCAGCTCGGGCCCG
TCTGCTCACAGAGATCCACGGGGGCGCCGGCGGGCCCAGCGGCAGGGCCGAAGCCGGTCGGGCCCTGGC
CAGTGGCACTGCAGCTTCCTCCACCGACCCTGGGGCCCCAGGGGGCCCGGGAGGGGCTGAGGGTCCCAT
GGCCAAGAAGCATGCTGGCGAGCGCGATAAGAAGCTGGCGGCCAAGAAAAAGACGGACAAGAAGCGGGC
GCTGCGGCGGCTGTAGTGGGCTCTCTTCCTCCTTCCACCGTGACCCCAACCTCTCCTGTCCCCTCCCTC
CAACTCTGTCTCTAAGTTATTTAAATTATGGCTGGGGTCGGGGAGGGTACAGGGGGCACTGGGACCTGG
ATTTGTTTTTCTAAATAAAGTTGGAAAAGCAAAAAAAAAAGCTT

Predicted Protein Sequence

MNSNVENLPPHIIRLVYKEVTTLTADPPDGIKVFPNEEDLTDLQVTIEGPEGTPYAGGLFRMKLLLGKD
FPASPPKGYFLTKIFHPNVGANGEICVNVLKRDWTAELGIRHVLLTIKCLLIHPNPESALNEEAGRLLL
ENYEEYAARARLLTEIHGGAGGPSGRAEAGRALASGTAASSTDPGAPGGPGGAEGPMAKKHAGERDKKL
AAKKKTDKKRALRRL

Fig. 1E

BR1-46

Nucleotide Sequence

GCCGGGGGCGCCCGCCGGTCGGGCCTGGGCGGCCGCCATGAAGCTGACGCGGAAGATGGTTCTGACCCG
AGCCAAGGCCTCGGAGCTGCACAGCGTGCGCAAGCTCAACTGCTGGGGCAGCCGCCTCACAGATATCTC
CATTTGCCAGGAGATGCCCAGCCTGGAGGTGATCACGCTCAGTGTCAACAGCATCTCCACCCTGGAGCC
TGTGAGCCGGTGCCAGCGCCTGAGTGAGCTGTACCTGCGGAGGAACCGCATCCCCAGCCTGGCTGAGCT
CTTCTACCTGAAGGGGCTGCCGCGTCTGCGGGTGCTGTGGCTGGCCGAGAACCCGTGCTGCGGCACCAG
CCCCCACCGCTACCGCATGACCGTGCTGCGCACCCTGCCGCGCCTACAGAAGCTGGACAACCAGGCTGT
GACGGAGGAGGAGCTGTCCCGTGCACTGAGTGAGGGAGAGGAGATCACTGCGGCCCCAGAGAGAGAGGG
CACAGGCCACGGCGGCCCCAAGCTATGCTGCACACTGAGCTCCCTCAGCTCCGCTGCTGAGACTGGCCG
GGACCCGCTGGACAGCGAGGAGGAGGCAACCAGCGGCGCCCAGGATGAACGTGGCCTGAAGCCGCCTTC
CCGGGGCCAGTTTCCTTCCCTCTCAGCCAGGGATGCCTCGAGCAGCCACAGGGGCAGGAACGTCCTGAC
TGCCATCCTGCTGCTGCTGCGGGAGCTGGATGCAGAGGGGCTGGAGGCCGTGCAGCAGACTGTGGGCAG
CCGGCTGCAGGCCCTGCGTGGGGAAGAGGTGCAGGAGCACGCCGAGTGACCGCAGGACCTGAACGCCGC
TCCAGCCTCCACGGGGACCCCAGCGTCTTCCCCAGCCCCCGGGAGCTGGAGGGTGGCTGCCATGGCCGC
AGCCCCGGCCCCACACAAAAGCCTCCCCGGTTTGCCACATCGGCCGAGGGCAGGAGTGGGTGTTAGGTA
CTGGCTAACCGGGGCGGTGGAGATGCCTGTCTACACCAGTCCTGTCCCCAGGACTCCCCTTCTGTGGTC
TGGAGGTTCTAGGCTGGCCTGGGCTCTTAAAGGGAGGATTTTGCAGGCTGTCCTCCCTAATAAAAGATT
TTCCCAAGGTTAAAAAAAAAAAGCTT

Predicted Protein Sequence

MKLTRKMVLTRAKASELHSVRKLNCWGSRLTDISICQEMPSLEVITLSVNSISTLEPVSRCQRLSELYL
RRNRIPSLAELFYLKGLPRLRVLWLAENPCCGTSPHRYRMTVLRTLPRLQKLDNQAVTEEELSRALSEG
EEITAAPEREGTGHGGPKLCCTLSSLSSAAETGRDPLDSEEEATSGAQDERGLKPPSRGQFPSLSARDA
SSSHRGRNVLTAILLLLRELDAEGLEAVQQTVGSRLQALRGEEVQEHAE

BR1-50

Nucleotide Sequence 5'

AGACGCTCGGTTCCCCGCCGTGCCTCCTCGCTGGCCGCGCTCCCTCCCGGTGCCGGCTTCTCTGAGTCA
CCAACCTGAGGCTGCCCCGGCCGCCTGCGCACCCGGCAGCACCATGACAGATCAGGCCTTTGTGACACT
AACCACAAACGATGCCTACGCCAAAGGTGCCCTGGTCCTGGGATCATCTCTGAAACAGCACAGGACCAC
CAGGAGGCTGGTCGTGCTCGCCACCCCTCAGGTCTCAGACTCCATGAGAAAAGTTTTAGAGACAGTCTT
TGATGAAGTCATCATGGTAGATGTCTTGGACAGTGGCGATTCTGCTCATCTAACCTTAATGAAGAGGCC
AGAGTTGGGTGTCACGCTGACAAAGCTCCACTGCTGGTCGCTTACACAGTATTCAAAATGTGTATTCAT
GGATGCAGATACTCTGGTCCTAGCAAATATTGATGATCTTTTTGACAGAGAAGAATTGTCAGCAGCACC
AGACCCAGGGTGGCCTGACTGCTTCAATTCCGGAGTCTTCGTTTATCAGCCTTCAGTTGAAACATACAA
TCAGCTGTTGCATCTTGCTTCTGAGCAAGGTA

Predicted Protein Sequence

MTDQAFVTLTTNDAYAKGALVLGSSLKQHRTTRRLVVLATPQVSDSMRKVLETVFDEVIMVDVLDSGDS
AHLTLMKRPELGVTLTKLHCWSLTQYSKCVFMDADTLVLANIDDLFDREELSAAPDPGWPDCFNSGVFV
YQPSVETYNQLLHLASEQG

Fig. 1F

BR1-51

Nucleotide Sequence 5'

AGGCAAAAAACAAATAGAAAAACTGGAACAGGAACTTAAAAGGTGTAAATCTGAGCTTGAAAGAAGCCA
ACAAGCTGCGCAGTCTGCAGATGTCTCTCTGAATCCATGCAATACACCACAAAAAATTTTTACAACTCC
ACTAACACCAAGTCAATATTATAGTGGTTCCAAGTATGAAGATCTAAAAGAAAAATATAATAAAGAGGT
TGAAGAACGAAAAAGATTAGAGGCAGAGGTTAAAGCCTTGCAGGCTAAAAAAGCAAGCCAGACTCTTCC
ACAAGCCACCATGAATCACCGCGACATTGCCCGGCATCAGGCTTCATCATCTGTGTTCTCATGGCAGCA
AGAGAAGACCCCAAGTCATCTTTCATCTAATTCTCAAAGAACTCCAATTAGGAGAGATTTCTCTGCATC
TTACTTTTCTGGGGAACAAGAGGTGACTCCAAGTCGATCAACTTTGCAAATAGGGAAAAGAGATGCTAA
TAGCAGTTTCTTTGACAATTCTAGCAGTCCTCATCTTTTGGATCAATTAAAAGCGCAGAATCAAGAGCT
AAGAAACAAGATTAATGAGTTGGAACTACGCCTGCAAGGACATGAAAAAGAAATGAAAGGCCAAGTGAA
TAAGTTTCAAGAACTCCAACTC

Predicted Protein Sequence

KKQIEKLEQELKRCKSELERSQQAAQSADVSLNPQNTPQKIFTTPLTPSQYYSGSKYEDLKEKYNKEVE
ERKRLEAEVKALQAKKASQTLPQATMNHRDIARHQASSSVFSWQQEKTPSHLSSNSQRTPIRRDFSASY
FSGEQEVTPSRSTLQIGKRDANSSFFDNSSSPHLLDQLKAQNQELRNKINELELRLQGHEKEMKGQVNK
FQELQL

BR1-58

Nucleotide Sequence

AAGAAAAAGAAGAAGAAGAAGCCTTTTATGTTAGATGAGGAAGGGGATACCCAAACAGAGGAAACCCAG
CCTTCAGAAACAAAAGAAGTGGAGCCAGAGCCAACTGAGGACAAGGATTTGGAAGCTGATGAAGAGGAC
ACTAGGAAAAAAGATGCTTCTGATGATCTAGATGACTTGAACTTCTTTAATCAAAAGAAAAAGAAGAAA
AAAACTAAAAAGATATTTGATATTGATGAAGCTGAAGAAGGTGTAAAGGATCTTAAGATTGAAAGTGAT
GTTCAAGAACCAACTGAACCAGAGGATGACCTTGACATTATGCTTGGCAATAAAAAGAAGAAAAAGAAG
AATGTTAAGTTCCCAGATGAGGATGAAATACTAGAGAAAGATGAAGCTCTAGAAGATGAAGACAACAAA
AAAGATGATGGTATCTCATTCAGTAATCAGACAGGCCCTGCTTGGGCAGGCTCAGAAAGAGACTACACA
TACGAGGAGCTGCTGAATCGAGTGTTCAACATCATGAGGGAAAAGAATCCAGATATGGTTGCTGGGGAG
AAAAGGAAATTTGTCATGAAACCTCCACAAGTCGTCCGAGTAGGAACCAAGAAAACTTCTTTTGTCAAC
TTTACAGATATCTGTAAACTATTACATCGTCAGCCCAAACATCTCCTTGCATTTTTGTTGGCTGAATTG
GGTACAAGTGGTTCTATAGATGGTAATAACCAACTTGTAATCAAAGGAAGATTCCAACAGAAACAGATA
GAAAATGTCTTGAGAAGATATATCAAGGAATATGTCACTTGTCACACATGCCGATCACCGGACACAATC
CTGCAGAAGGACACACGACTCTATTTCCTACAGTGCGAAACTTGTCATTCTAGATGTTCTGTTGCCAGT
ATCAAAACCGGCTTCCAGGCTGTCACGGGCAAGCGAGCACAGCTCCGTGCCAAAGCTAACTAATTTGCT
AATCACTGATTTTGCAAAGCTTGTTGTGGAGATGTGGCTGGACAGGTTTGCCATCAGAGTGGATATACC
GTTGTATTAAAAACAAGATAAAAAAGCTGCCAAGATTTTTGGCGAGTGGTTGGTCTGAAGTCCTTGCAA
GACGCTGATGCTCAAGCTGTTGACATACTCATTGCCTACTTTAACACCTGTCAGAGAAACGTGATATGG
GGTAAGGAGGTGCTTTTTTTAAAATCGTTCATAGACTTCTGTAAAATGCAAGATAAATTAAAGTTATTAT
AACAGTGAAAGCTT

Fig. 1G

Predicted Protein Sequence

KKKKKKKPFMLDEEGDTQTEETQPSETKEVEPEPTEDKDLEADEEDTRKKDASDDLDDLNFFNQKKKKK
KTKKIFDIDEAEEGVKDLKIESDVQEPTEPEDDLDIMLGNKKKKKKNVKFPDEDEILEKDEALEDEDNK
KDDGISFSNQTGPAWAGSERDYTYEELLNRVFNIMREKNPDMVAGEKRKFVMKPPQVVRVGTKKTSFVN
FTDIC

BR1-92

Nucleotide Sequence

CAAGTACACACTCAAGATTGTCGAAGAGGAAACCAAAAAAGAGCTGGAAGAGAACAAGCGATCCCTGGC
TGCATTGGATGCACTCAATACTGATGATGAAAATGATGAGGAGGAATATGAGGCATGGAAAGTTCGAGA
GCTAAAAAGAATCAAGAGGGACAGAGAAGATCGAGAAGCGCTTGAGAAGGAGAAAGCAGAAATTGAACG
CATGCGAAACCTGACTGAGGAAGAGAGGAGAGCTGAACTTCGGGCAAACGGCAAAGTCATTACCAACAA
AGCTGTTAAGGGCAAATACAAGTTCTTACAGAAGTATTATCACCGGGGTGCCTTCTTCATGGATGAGGA
TGAAGAAGTATACAAGAGAGATTTCAGCGCTCCTACCCTGGAGGATCATTTCAATAAAACCATTCTTCC
TAAAGTCATGCAGGTCAAGAAC

BR1-97

Nucleotide Sequence

GGCGAGGACGTGAAGCCTCCCGAAGACGCACAGAAGGAAGGAGCCAGCTCCCAGCCCACTCATCGCAGG
GCTCATGATTTTTTACAAATTATGTTTTAATTCCAAGTGTTTCTGTTTCAAGGAAGGATGAATAAGTTT
TATTGAAAATGTGGTAACTTTATTTAAAATGATTTTTAACATTATGAGAGACTGCTCAGATTCTAAGTT
GTTGGCCTTGTGTGTGTGTTTTTTTTTAAGTTCTCATCATTATTACATAGACTGTGAAGTATCTTTACT
GGAAATGAGCCCAAGCACACATGCATGGCATTTGTTCCACAGGAGGGCATCCCTGGGGATGTGGCTGGA
GCATGAGCCAGCTCTGTCCCAGGATGGTCCCAGCGGATGCTGCCAGGGGCAGTGAAGTGTTTAGGTGAA
GGACAAGTAGGTAAGAGGACGCCTTCAGGCACCACAGATAAGCCTGAAACAGCCTCTCCAAGGGTTTTC
ACCTTAGCAACAATGGGAGCTGTGGGAGTGATTTTGGCCACACTGTCAACATTTGTTAGAACCAGTCTT
TTGAAAGAAAAGTATTTCCAACTTGTCACTTGCCAGTCACTCCGTTTTGCAAAAGGTGGCCCTTCACTG
TCCATTCCAAATAGCCCACACGTGCTCTCTGCTGGATTCTAAATTATGTGAATTTTGCCATATTAAATC
TTCCTCATTTATACTATTATTTGTTACGTTCAATCAGAATCCCCGAAACCTCCTATAAAGCTTAGCTGC
CCCTTCTGAGGATGCTGAGAACGGTGTCTTTCTTTATAAATGCAAATGGCTACCGTTTTACAATAAAAT
TTTGCATGTGCAAAAAAAAAAGCTT

BR1-98

Nucleotide Sequence

ATCCCCTCCCCGCGCGTTCCTTCGCACACTGTGATTTTGCCCTCCTGCCCACGCAGACCTGCAGCGGGC
AAAGAGCTCCCGAGGAAGCACAGCTTGGGTCAGGTTCTTGCCTTTCTTAATTTTAGGGACAGCTACCGG
AAGGAGGGGAACAAGGAGTTCTCTTCCGCAGCCCCTTTCCCCACGCCCACCCCCAGTCTCCAGGGACCC
TTGCCTGCCTCCTAGGCTGGAAGCCATGGTCCCGAAGTGTAGGGCAAGGGTGCCTCAGGACCTTTTGGT
CTTCAGCCTCCCTCAGCCCCCAGGATCTGGGTTAGGTGGCCGCTCCTCCCTGCTCCTCATGGGAAGATG
TCTCAGAGCCTTCCATGACCTCCCCTCCCCAGCCCAATGCCAAGTGGACTTGGAGCTGCACAAAGTCAG
CAGGGACCACTAAATCTCCAAGACCTGGTGTGCGGAGGCAGGAGCATGTATGTCTGCAGGTGTCTGACA
CGCAAGTGTGTGAGTGTGAGTGTGAGAGATGGGGCGGGGGTGTGTCTGTAGGTGTCTCTGGGCCTGTGT
GTGGGTGG

Fig. 1H

BR1-107

Nucleotide Sequence

CAGCATCTCGCCTCGGCCGAGGCCCCCCGGGCGCCCCGTCAGCGGTCACGGGATGGATAGCCGGCCTGC
CATGGCCATCTTTGAACTCCTGGACTATATTGTGAACGAGCCACCTCCTAAGCTGCCCAACGGTGTGTT
CACCCCCGACTTCCAGGAGTTTGTCAATAAATGCCTCATCAAGAACCCAGCGGAGCGGGCGGACCTGAA
GATGCTCACAAACCACACCTTCATCAAGCGGTCCGAGGTGGAAGAAGTGGATTTTGCCGGCTGGTTGTG
TAAAACCCTGCGGCTGAACCAGCCCGGCACACCCACGCGCACCGCCGTGTGACAGTGGCCGGGCTCCCT
GCGTCCCGCTGGTGACCTGCCCACCGTCCCTGTCCATGCCCCGCCCTTCCAGCTGAGGACAGGCTGGCG
CCTCCACCCACCCTCCTGCCTCACCCCTGCGGAGAGCACCGTGGCGGGGCGACAGCGCATGCAGGAACG
GGGGTCTCCTCTCCTGCCCGTCCTGGCCGGGGTGCCTCTGGGGACGGGCGACGCTGCTGTGTGTGGTCT
CAGAGGCTCTGCTTCCTTAGGTTACAAAACAAAACAGGGAGAGAAAAAGCAAAAAAAAAAAAAAAAAAA
AAAGCTT

BR1-110

Nucleotide Sequence

TGCGGGAATAGGTGCAGCGGGCCCTTGGCGGGGGACTCTGAGGGAGGAGCTGGGGACGGCGACCCTAGG
AGAGTTCTTTGGGGTGACTTTCAAGATGGACTCTACTCTAACAGCAAGTGAAATCCGGCAGCGATTTAT
AGATTTCTTCAAGAGGAACGAGCATACGTATGTTCACTCGTCTGCCACCATCCCATTGGATGACCCCAC
TTTGCTCTTTGCCAATGCAGGCATGAACCAGTTTAAACCCATTTTCCTGAACACAATTGACCCATCTCA
CCCCATGGCAAAGCTGAGCAGAGCTGCCAATACCCAGAAGTGCATCCGGGCTGGGGGCAAACATAATGA
CCTGGACGATGTGGGCAAGGATGTCTATCATCACACCTTCTTCGAGATGCTGGGCTCTTGGTCTTTTGG
AGATTACTTTAAGGAATTGGCATGTAAGATGGCTCTGGAACTCCTCACCCAAGAGTTTGGCATTCCCAT
TGAAAGACTTTATGTTACTTACTTTGGCGGGGATGAAGCAGCTGGCTTAGAAGCAGATCTGGAATGCAA
ACAGATCTGGCAAAATTTGGGGCTGGATGACACCAAAATCCTCCCAGGCAACATGAAGGATAACTTCTG
GGAGATGGGTGACACG

BR1-113

Nucleotide Sequence

GGAGTAGGTGCGGGTGAAGATGGCGGCAGCCGAGGCCGCGAACTGCATCATGGAGGTGTCCTGTGGCCA
GGCGGAAAGCAGTGAGAAGCCCAACGCTGAGGACATGACATCCAAAGATTACTACTTTGACTCCTACGC
ACACTTTGGCATCCACGAGGAGATGCTGAAGGACGAGGTGCGCACCCTCACTTACCGCAACTCCATGTT
TCATAACCGGCACCTCTTCAAGGACAAGGTGGTGCTGGACGTCGGCTCGGGCACCGGCATCCTCTGCAT
GTTTGCTGCCAAGGCCGGGGCCCGCAAGGTCATCGGGATCGAGTGTTCCAGTATCTCTGATTATGCGGT
GAAGATCGTCAAAGCCAACAAGTTAGACCACGTGGTGACCATCATCAAGGGGAAGGTGGAGGAGGTGGA
GCTCCCAGTGGAGAAGGTGGACATCATCATCAGCGAGTGGATGGGCTACTGCCTCTTCTACGAGTCCAT
GCTCAACACCGTGCTCTATGCCCGGGACAAGTGGCTGGCGCCCGATGGCCTCATCTTCCCAGACCGGGC
CACGCTGTATGTGACGGCCATCGAGG

Fig. 1I

BR1-119

Nucleotide Sequences

CGGGACTTCTCTGAATGGTCGGTGCACTCAGGTGATTCCTCCCCTGGGCTCCCAGAGGCAGCAAACCCA
TTATACTGGAACCTAGGCCCTTCCTGAGTTTCCCCTCCACACAGCTAGGAGCCCATGCCCGGCCTGATC
TCAGCCCGAGGACAGCCCCTCCTTGAGGTCCTTCCTCCCCAAGCCCACCTGGGTGCCCTCTTTCTCCCT
GAGGCTCCACTTGGTCTCTCCGCGCAGCCTGCCCTGTGGCCCACCCTGGCCGCTCTGGCTCTGCTGAGC
AGCGTCGCAGAGGCCTCCCTGGGCTCCGCGCCCCGCAGCCCTGCCCCCCGCGAAGGCCCCCCGCCTGTC
CTGGCGTCCCCCGCCGGCCACCTGCCGGGTAGGTGAGAGGGCGAGGGGGCGGGGCGGGGCTGGCCCGGG
ACACCGCGCGTGACTGGGTCTCATTCCAGGGGGACGCACGGCCCGCTGGTGCAGTGGAAGAGCCCGGCG
GCCGCCGCCGCAGCCTTCTCGGCCCGCGCCCCCGCCGCCTGCACCCCCATC

Predicted Protein Sequence

MPGLISARGQPLLEVLPPQAHLGALFLPEAPLGLSAQPALWPTLAALALLSSVAEASLGSAPRSPAPRE
GPPPVLASPAGHL

BR1-123

Nucleotide Sequence

GCCGCCGGGGCCTCTCCAGGGCCGCAGCGGCAGCAGTTGGGCCCCCCGCCCCGGCCGGCGGACCGAAGA
ACGCAGGAAGGGGGCCGGGGGGACCCGCCCCCGGCCGGCCGCAGCCATGAACTCCAACGTGGAGAACCT
ACCCCCGCACATCATCCGCCTGGTGTACAAGGAGGTGACGACACTGACCGCAGACCCACCCGATGGCAT
CAAGGTCTTTCCCAACGAGGAGGACCTCACCGACCTCCAGGTCACCATCGAGGGCCCTGAGGGGACCCC
ATATGCTGGAGGTCTGTTCCGCATGAAACTCCTGCTGGGGAAGGACTTCCCTGCCTCCCCACCCAAGGG
CTACTTCCTGACCAAGATCTTCCACCCGAACGTGGGCGCCAATGGCGAGATCTGCGTCAACGTGCTCAA
GAGGGACTGGACGGCTGAGCTGGGCATCCGACACGTACTGCTGACCATCAAGTGCCTGCTGATCCACCC
TAACCCCGAGTCTGCACTCAACGAGGAGGCGGGCCGCCTGCTC

BR1-124

Nucleotide Sequence

CATGCGCCCGCGGCTGCGCTACAGCCCCTACTCCATCCCGGTGCCGGTCCCGGACGGCAGCAGTCTGCT
CACCACCGCCCTGCCCTCCATGGCGGCGGCCGCGGGGCCCCTGGACGGCAAAGTCGCCGCCCTGGCCGC
CAGCCCGGCCTCGGTGGCAGTGGACTCGGGCTCTGAACTCAACAGCCGCTCCTCCACGCTCTCCTCCAG
CTCCATGTCCTTGTCGCCCAAACTCTGCGCGGAGAAAGAGGCGGCCACCAGCGAACTGCAGAGCATCCA
GCGGTTGGTTAGCGGCTTGGAAGCCAAGCCGGACAGGTCCCGCAGCGCGTCCCCGTAGACCCGTCCCAG
ACACGTCTTTTCATTCCAGTCCAGTTCAGGCTGCCGTGCACTTTGTCGGATATAAAATAAACCACGGGC
CCGCCATGGCGTTAGCCCTTCCTTTTGCAGTTGCGTCTGGGAAGGGGCCCCGGACTCCCTCGAGAGAAT
GTGCTAGAGACAGCCCCTGTCTTCTTGGCGTGGTTTATATGTCCGGGATCTGGATCAGATTCTGGGGGC
TCAGAAACGTCGGTTGCATTGAGCTACTGGGGGTAGGAGTTCCAACATTTATGTCCAGAGCAACTTCCA
GCAAGGCTGGTCTGGGTCTCTGCCCACCAGGCGGGGAGGTGTTCAAAGACATCTCCCTCAGTGCGGATT
TATATATATATTTTTCCTTCACTGTGTCAAGTGGAAACAAAAACAAAATCTTTCAAAAAAAAAAAAGCT
T

Fig. 1J

Predicted Protein Sequence
MRPRLRYSPYSIPVPVPDGSSLLTTALPSMAAAAGPLDGKVAALAASPASVAVDSGSELNSRSSTLSSS
SMSLSPKLCAEKEAATSELQSIQRLVSGLEAKPDRSRSAS

BR1-189

Nucleotide Sequence 5'
GAGGAAAGTGCCACACGCTAGAGGAGGTCTGGTCTTGTTGGATTGAACTGCTTCACTATTTGGATCTTG
AAACTACCTGGTTAAACACTTTGGAAGAGCGGATGAAGAGCACAGAGGTCCTGCCTGAGAAGACGGATG
CTGTCAACGAAGCCCTGGAGTCTCTGGAATCTGTTCTGCGCCACCCGGCAGATAATCGCACCCAGATTC
GAGAGCTTGGCCAGACTCTGATTGATGGGGGGATCCTGGATGATATAATCAGTGAGAAACTGGAGGCTT
TCAACAGCCGATATGAAGATCTAAGTCACCTGGCAGAGAGCAAGCAGATTTCTTTGGAAAAGCAACTCC
AGGTGCTGCGGGAAACTGACCAGATGCTTCAAGTCTTGCAAGAGAGCTTGGGGGAGCTGGACAAACAGC
TCACCACATACCTGACTGACAGGATAGATGCTTTCCAAGTTCCACAGGAAGCTCAGAAAATCCAAGCAG
AGATCTCAGCCCATGAGCTAACCCTAGAGGAGTTGAGAAGAAATA

Predicted Protein Sequence
GKCHTLEEVWSCWIELLHYLDLETTWLNTLEERMKSTEVLPEKTDAVNEALESLESVLRHPADNRTQIR
ELGQTLIDGGILDDIISEKLEAFNSRYEDLSHLAESKQISLEKQLQVLRETDQMLQVLQESLGELDKQL
TTYLTDRIDAFQVPQEAQKIQAEISAHELTLEELRRN

BR1-194

Nucleotide Sequence 5'
GGAAGAAGAGGCGAGAACGACCCCCGGACCGACCAAAGCCCGCGCGCCGCTGCATCCCGCGTCCAGCAC
CTACGTCCCGCTGCCGTCGCCGCCGCCACCATGCCCAAGAGAAAGGCTGAAGGGGATGCTAAGGGAGAT
AAAGCAAAGGTGAAGGACGAACCACAGAGAAGATCCGCGAGGTTGTCTGCTAAACCTGCTCCTCCAAAG
CCAGAGCCCAAGCCTAAAAAGGCCCCTGCAAAGAAGGGAGAGAAGGTACCCAAAGGGAAAAAGGGAAAA
GCTGATGCTGGCAAGGAGGGGAATAACCCTGCAGAAAATGGAGATGCCAAAACAGACCAGGCACAGAAA
GCTGAAGGTGCTGGAGATGCCAAGTGAAGTGTGTGCATTTTTGATAACTGTGTACTTCTGGTGACTGTA
CAGTTTGAAATACTATTTTTTATCAAGTTTTATAAAAATGCAGAATTTTGTTTTACTTTTTTTTTTTTT
TAAAAGCTATGTTGTTAGCACACAGAACACTTCATTGTTGTTT

Predicted Protein Sequence
MPKRKAEGDAKGDKAKVKDEPQRRSARLSAKPAPPKPEPKPKKAPAKKGEKVPKGKKGKADAGKEGNNP
AENGDAKTDQAQKAEGAGDAK

Fig. 1K

BR1-201

Nucleotide Sequence

CCGTCGCGTACCTAGGATGCCGCGTGGAAGCCGAAGCCGCACCTCCCGCATGGCCCCTCCGGCCAGCCG
GGCCCCTCAGATGAGAGCTGCACCCAGGCCAGCACCAGTCGCTCAGCCACCAGCAGCGGCACCCCCATC
TGCAGTTGGCTCTTCTGCTGCTGCGCCCCGGCAGCCAGGTCTGATGGCCCAGATGGCAACCACTGCAGC
TGGCGTGGCTGTGGGCTCTGCTGTGGGGCACACATTGGGTCACGCCATTACTGGGGGCTTCAGTGGAGG
AAGTAATGCTGAGCCTGCGAGGCCTGACATCACTTACCAGGAGCCTCAGGGAACCCAGCCAGCACAGCA
GCAGCAGCCTTGCCTCTATGAGATCAAACAGTTTCTGGAGTGTGCCCAGAACCAGGGTGACATCAAGCT
CTGTGAGGGTTTCAATGAGGTGCTGAAACAGTGCCGACTTGCAAACGGATTGGCCTAATGAAGAAGTTC
AACCTGGAGAGATGGAAAATCAGCTCTCATAACTAAGTTAATTTAGTATAAAAATAGAATTGATAGTGA
GGGTATAAAGTGTAACCATCAGTTAAACCTCTCCTGTCATTCCTAGCTTCCTTGCTTCAGAATTGAAAT
GGAAGTGGGGGTGTCCCTACTCTGTAGAATCTGGGACTGGGCAAATGTTTGTGTGGCCTCCTTAAACTA
GCTGTTATGTTATGATTTTATTCTTTGTGAGTTAATTAGAATAAAGTCATTTTCTTCCAAGGAAAAAAA
AAAAAAAAAAGCTT

Predicted Protein Sequence

MPRGSRSRTSRMAPPASRAPQMRAAPRPAPVAQPPAAAPPSAVGSSAAAPRQPGLMAQMATTAAGVAVG
SAVGHTLGHAITGGFSGGSNAEPARPDITYQEPQGTQPAQQQQPCLYEIKQFLECAQNQGDIKLCEGFN
EVLKQCRLANGLA

BR1-210

Nucleotide Sequence

TTTATTAACACCACCCGAGCAGGTCCAGGGACATTATCCGTCACCATCGAAGGCCCATCCAAGGTTAAA
ATGGA
TTGCCAGGAAACACCTGAAGGGTACAAAGTCATGTACACCCCCATGGCTCCTGGTAACTACCTGATCAG
CGTCAAATACGGTGGGCCCAACCACATCGTGGGCAGTCCCTTCAAGGCCAAGGTGACAGGCCAGCGTCT
AGTTAGCCCTGGCTCAGCCAACGAGACCTCATCCATCCTGGTGGAGTCAGTGACCAGGTCGTCTACAGA
GACCTGCTATAGCGCCATTCCCAAGGCATCCTCGGACGCCAGCAAGGTGACCTCTAAGGGGGCAGGGCT
CTCAAAGGCCTTTGTGGGCCAGAAGAGTTCCTTCCTGGTGGACTGCAGCAAAGCTGGCTCCAACATGCT
GCTGATCGGGGTCCATGGGCCCACCACCCCCTGCGAGGAGGTCTCCATGAAGCATGTAGGCAACCAGCA
ATACAACGTCACATACGTCGTCAAGGAGAGGGGCGATTATGTGCTGGCTGTGAAGTGGGGGGAGGAACA
CATCCCTGGCAGCCCTTTTCATGTCACAGTGCCTTAAAACAGTTTTCTCAAAAAAAAAAGCTT

BR1-213

Nucleotide Sequence

GTTCAGTGTCTGCAACGGGCCACTCCAGAACAGTACCAGATCCTGAAGGAAAATTACGGGCAGAAGGAG
GCTGAGAAAGTGGCCCGGGTGAAGGCGCTATATGAGGAGCTGGATCTGCCAGCAGTGTTCTTGCAATAT
GAGGAAGACAGTTACAGCCACATTATGGCTCTCATTGAACAGTACGCAGCACCCCTGCCCCCAGCCGTC
TTTCTGGGGCTTGCGCGCAAAATCTACAAGCGGAGAAAGTGACCTAGAGATTGCAAGGGCGGGGAGAGG
AGGCTCTCAATAAATAATCGTGTAACCTTAAAAAAAAAAAAAGCTT

Fig. 1L

BR1-219

Nucleotide Sequence 5'

CCGACCCGGAGACCCGGTTCTCTTTGATCCGCCAGGTCCTCAATGTAGGACCCCGCGTGCCCATTGGGC
CTAATCCCGTGATCACTGAACAGCTACCCCCCTCCCAACCCGTGCAGATCATGCTCCCCAGGCCTCCTC
ATCCTCCTCCTTCAGGCGCGGCCTCTATGGTGCCTGGGGCTCCCCCGCCTTCTCAACAACCTGGGACGG
GGGACAGGCTGCTAAACCTAGTAAAAGGAGCCTATCAAGCACTCAACCTCACCAGTCCCGACAGAACCC
AAGAGTGCTGGCTGTGTCTGGTATCGGGACCCCCCTACTACGAAGGGGTTGCCGTCCTAGGTACCTACT
CCAACCATACCTCTGCCCCAGCTAACTGCTCCGTGGCCTCCCAACACAAGCTGACCCTGTCCGAAGTGA
CCCGGCAGGGACTCTGCGTAGGAGCAGTTCCCAAAACCCATCAGGCCCTGTGTAATACCACCCAGAAGG
CGAGCGACGGGTCCTACTATCTGGCTGCTCCCGCCGGGACCATCTGGGCTTGC

BR1-220

Nucleotide Sequence

GGAGGCCGAGGCGCCGGAGCAGGAGGAGGCCGGCCGGAGGCGGCATGAGACGAGCGTGGGGGCCGCGGC
TGCTCGGGGCCGCGCTGGTTGCCCATTGACAGCGGCGTCTGCAGCTCGCTTCAAGATGGCCGCTTGGCT
CGCATTCATTTTCTGCTGAACGACTTTTAACTTTCATTGTCTTTTCCGCCCGCTTCGATCGCCTCGCGC
CGGCTGCTCTTTCCGGGATTTTTTATCAAGCAGAAATGCATCGAACAACGAGAATCAAGATCACTGAGC
TAAATCCCCACCTGATGTGTGTGCTTTGTGGAGGGTACTTCATTGATGCCACAACCATAATAGAATGTC
TACATTCCTTCTGTAAAACGTGTATTGTTCGTTACCTGGAGACCAGCAAGTATTGTCCTATTTGTGATG
TCCAAGTTCACAAGACCAGACCACTACTGAATATAAGGTCAGATAAAACTCTCCAAGATATTGTATACA
AATTAGTTCCAGGGCTTTTCAAAAATGAAATGAAGAGAAGAAGGGATTTTTATGCAGCTCATCCTTCTG
CTGATGCTGCCAATGGCTCTAATGAAGATAGAGGAGAGGTTGCAGATGAAGATAAGAGAATTATAACTG
ATGATGAGATAATAAGCTT

BR1-221

Nucleotide Sequence 3'

AATGGAGATGCTGACCAAAGACAAAGCTGACAAAGATGAACAAATCAGAAAAATAAAATCTAGGCACAG
TGATGAATTAACCTCACTGTTGGGATATTTTCCCACAAAAAACAGCTTGAAGACTGGCTACATAGTAAA
TCAAAAGAAATTAATCAGACCAGGGACAGACTTGCCAAATTGAACAAGGAACTAGCTTCATCTGAGCAG
AATAAAAATCATATAAATAATGAACTAAAAAGAAGGAAGAGCAGTTGTCCAGTTACGAAGACAAGCTGT
TTGATGTTTGTGGTAGCCAGGATTTTGAAAGTGATTTAGACAGGCTTAAAGAGGAAATTGAAAAATCAT
CAAAACAGCGAGCCATGCTGGCTGGAGCCACAGCAGTTTACTCCCAGTTCATTACTCAGCTAACAGACG
AAAACCAGTCATGTTGCCCCGTTTGTCAGAGAGTTTTTCAGACAGAGGCTGAGTTACAAGAAGTCATCA
GTGATTTGCAGTCTAAACTGCGACTTGCTCCAGATAAACTCAAGTCAACAGAATCAGAGCTAAAAAAAAA
AAAAAAAGCTT

Fig. 1M

BR1-222

Nucleotide Sequence 5'

AGGAAGGCAGGGGGGAAGGGACAGTCGGCCGCAGACCGCGCTGGGTTGCCGCTGCCGCTGCCGCCATCG
TGCCAGCCCCTCGGGTCTCCGTGAGGCCGGGTGACGCTCCAGAATGGGAGACAAGCCAATTTGGGAGCA
GATTGGATCCAGCTTCATTCAACATTACTACCAGTTATTTGATAATGATAGAACCCAACTAGGCGCAAT
TTACATTGACGCGTCATGCCTTACGTGGGAAGGACAACAGTTCCAGGGGAAAGCTGCCATTGTGGAGAA
GTTGTCTAGCCTTCCGTTCCAGAAAATTCAGCACAGCAT

BR1-224

Nucleotide Sequence

CAGCTTTTATAGTAAAGATACCTCTTTACGGACTCCACTTATGACTCCCTAAAGCCCATGTCGAAGCCC
CCATCGCTGGGTCAATAGTACTTGCCGCAGTACTCTTAAAACTAGGCGGCTATGGTATAATACGCCTCA
CACTCATTCTCAACCCCCTGACAAAACACATAGCCTACCCCTTCCTTGTACTATCCCTATGAGGCATAA
TTATAACAAGCTCCATCTGCCTACGACAAACAGACCTAAAATCGCTCATTGCATACTCTTCAATCAGCC
ACATAGCCCTCGTAGTAACAGCCATTCTCATCCAAACCCCCTGAAGCTTCACCGGCGCAGTCATTCTCA
TAATCGCCCACGGGCTTACATCCTCATTACTATTCTGCCTAGCAAACTCAAACTACGAACGCACTCACA
GTCGCATCATAATCCTCTCTCAAGGACTTCAAACTCTACTCCCACTAATAGCTTTTTGATGACTTCTAG
CAAGCCTCGCTAACCTCGCCTTACCCCCCACTATTAACCTACTGGGAGAACTCTCTGTGCTAGTAACCA
CGTTCTCCTGATCAAATATCACTCTCCTACTTACAGGACTCAACATACTAGTCACAGCCCTATACTCCC
TCTACATATTTACCACAACACAATGGGGCTCACTCACCCACCACATTAACAACATAAAACCCTCATTCA
CACGAGAAAACACCCTCATGTTCATACACCTATCCCCCATTCTCCTCCTATCCCTCAACCCCGACATCA
TTACCGGGTTTTCCTCTTAAAAAAAAAAAGCTT

BR1-232

Nucleotide Sequence

GGCCGCGGCGGGGGACAGCGACGAGGAGTCGCGGGCCGACGACAAGGGCGTCATGGACTACTACCTGAA
GTACTTCAGCGGCGCCCACGACGGCGACGTCTACCCGGCCTGGTCGCAGAAGGTGGAGAAGAAGATCCG
AGCCAAGGCCTTCCAGAAGTGCCCCATCTGCGAGAAGGTCATCCAGGGCGCCGGCAAGCTGCCGCGACA
CATCCGCACCCACACGGGCGAGAAGCCCTACGAGTGCAACATCTGCAAGGTCCGCTTCACCAGGCAGGA
CAAGCTGAAGGTGCACATGCGGAAGCACACGGGCGAGAAGCCGTACCTGTGCCAGCAGTGCGGCGCCGC
CTTTGCCCACAACTACGACCTGAAGAACCAC

Predicted Protein Sequence 1

GRGGGQRRGVAGRRQGRHGLLPEVLQRRPRRRRLPGLVAEGGEEDPSQGLPEVPHLREGHPGRRQAAAT
HPHPHGREALRVQHLQGPLHQAGQAEGAHAEAHGREAVPVPAVRRRLCPQLRPEE

Predicted Protein Sequence 2

AAAGDSDEESRADDKGVMDYYLKYFSGAHDGDVYPAWSQKVEKKIRAKAFQKCPICEKVIQGAGKLPRH
IRTHTGEKPYECNICKVRFTRQDKLKVHMRKHTGEKPYLCQQCGAAFAHNYDLKNH

Fig. 1N

BR1-234

Nucleotide Sequence

CTGGCGCAGCCAAGGCCGAAGCTCTGGCTGAACCCTGTGCTGGTGTCCTGACCACCCTCCCCTCTCTTG
CACCCGCCTCTCCCGTCAGGGCCCAAGTCCCTGTTTTCTGAGCCCGGGCTGCCTGGGCTGTTGGCACTC
ACAGACCTGGAGCCCCTGGGTGGGTGGTGGGGAGGGGCGC

BR1-237

Nucleotide Sequence

CCGAGTCGCGCGGAGGCGGAGGCTTGGGTGCGTTCAAGATTCAACTTCACCCGTAACCCACCGCCATGG
CCGAGGAAGGCATTGCTGCTGGAGGTGTAATGGACGTTAATACTGCTTTACAAGAGGTTCTGAAGACTG
CCCTCATCCACGATGGCCTAGCACGTGGAATTCGCGAAGCTGCCAAAGCCTTAGACAAGCGCCAAGCCC
ATCTTTGTGTGCTTGCATCCAACTGTGATGAGCCTATGTATGTCAAGTTGGTGGAGGCCCTTTGTGCTG
AACACCAAATCAACCTAATTAAGGTTGATGACAACAAGAAACTAGGAGAATGGGTAGGCCTTTGTAAAA
TTGACAGAGAGGGGAAACCCCGTAAAGTGGTTGGTTGCAGTTGTGTAGTAGTTAAGGACTATGGCAAGG
AGTCTCAGGCCAAGGATGTCATTGAAGAGTATTTCAAATGCAAGAAATGAAGAAATAAATCTTTGGCTC
ACAAAAAAAAAAAAAAAGGCTT

Fig. 10

BR1-16

Nucleotide Sequence

TGGGAACAATGTAGAGTATCAAGATAAGCTGGAGTGCAGGGTATGTATAGGCAAGTCATGGGTTTAGAA
GAAAAGCAGAT
GGCAAGACCTGCAGGACCAGATGCCAGGCTGCCTTGGTTTGAGAGGCAACGAAGGGTTAGAAGAATAGC
CTGGCTTGGCA
GTGACACGTCGGAGCTGTGCTTTAGAAATAAATGGATCCAGCAGCAGCAGGTGACCTGGTCTGGAGCTG
GAGAAGGGACC
AGCTGGTGGCGTCTGCACTCACTCAGGGCAGAAGTAGGGAATGTATAACCACAGCTGGCAGCACAGTGG
GTAGAGCGTAG
ACAGCTCCACAGATATTTGTTAAAGTGAGTTAGGTTTCTTATGGGCAGGAATTAGATCTATTTCTTTTT
TAAAAAAAAAA
AAGCTT

BR1-45

Nucleotide Sequence

ACACGGTGGGGCTCCCTGGCTGCGCCCGGCCCGGCAGCTACGGGCCCAGCGCCTGGTGGCGGCGCTGAG
GGGTCGCCGAGAGGGGCCCGGCGGCGTCTGCGGGGGCCGCTCCCTCGGTGGGCCGCGGGCGAGGCATGA
GCGCGGGCTCCCCCTGCCTCCGGAGCGCCGGCGGGGGACCGGGGGGCAGGAGATGTGCCTGTCCTTAGC
GGCCCAGGAAGCAGCATGCACCCCGATGCGACCGACAGTGGCGGCGCCGGCCCCAGCCCCGCGCGGGCC
GCAGGCGCCGGCGGCCGTCCTGTCTCGGGCTTCAGGGGCGAGCGGCGGCCGGAGTCCCCGGGGGACGCG
GAGGCAGCAGCAGCGGCGGCGCCGGGGGCCCCGGGCGGCCGGAGCTGGTGGAAGCCCGTGGCGGTGGCC
GCACTCGCCGCCGTGGCCCTCTCCTTCCTGGGGCCCGGCAGCGGGGAGGCGGCGGGGGCCGCGGGGCTG
AGCTCCGTCCTGTTCAGGCTCAGCCTGTACCTGAGCTGCGCGGCGGCCGCCTTCCTGCTGGGGATCCTG
TTTGCCCTCGTCTGCCGGAGCCCGCGCGCCCAGCCGCCCGACTTCGCCGCCGCCTGGAGCCGGCTGGCC
GCGACCTCAGCCGCCCGCCGCCCGCCGGGGAGTCCTGTGTATGGAAACTCACATGAGTCAGCTCAGTCT
AGAAGGGTAGTAATTTCTCATAATATGGATAAAGCTCTGAAAGAAGTGTTCGACTACAGTTATAGAGAT
TACATTCTGTCCTGGTATGGAAACCTCAGCAGAGATGAGGGACAACTTTACCATCTGCTCTTGGAAGAC
TTTTGGGAAATTGCCAGACAGCTGCACCACAGACTGAGTCACGTGGATGTGGTTAAAGTTGTCTGCAAT
GATGTTGTGAGGACTTTACTCACTCATTTCTGTGACCTGAAAGCTGCCAATGCCAGACATGAAGAACAG
CCAAGACCTTTTGTGTTGCACGCATGCTTGAGGAACTCAGATGATGAAGTAAGATTTCTACAAACGTGT
TCTCGGGTTCTGGTGTTTTGTCTCCTCCCCTCAAAGGATGTGCAGTCTCTCAGCTTACGTATAATGCTT
GCAGAAATTCTCACAACAAAAGTCTTGAAGCCGGTAGTGGAGTTACTGAGTAATCCAGATTACATTAAC
CAAATGCTGCTTGCCCAGCTGGCGTACAGAGAGCAAATGAATGAGCATCACAAGAGAGCCTACACCTAT
GCCCCCTCTTACGAGGACTTCATCAAGCTCATTAACAGCAACTCTGATGTGGAGTTCTTGAAGCAACTA
AGGTATCAAATTGTAGTGGAAATAATCCAGGCGACTACAATTAGCAGCTTTCCCCAACTGAAGAGGCAC
AAAGGTAAAGAAACTGCGGCAATGAAAGCTGATCTCCTGAGGGCCAGGAACATGAAGAGGTACATCAAC
CAACTGACTGTGGCAAAGAAGCAGTGTGAGAAGAGAATCCGAATCCTGGGAGGCCCTGCCTATGACCAG
CAAGAGGATGGGGCCCTGGATGAGGGGGAAGGGCCTCAAAGCCAGAAGATGGAGTCTCACTCTGTCGCC
CAGGCTGGCATGCAGTGAGCCGAGTTCGCGCCGCTGCCCTCAAGCCTGGCGACAGAGTGAGACTCTGTC
TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCTTG

Fig. 2A

Predicted Protein Sequence

MHPDATDSGGAGPSPARAAGAGGRPVSGFRGERRPESPGDAEAAAAAAPGAPGGRSWWKPVAVAALAAV
ALSFLGPGSGEAAGAAGLSSVLFRLSLYLSCAAAAFLLGILFALVCRSPRAQPPDFAAAWSRLAATSAA
RRPPGSPVYGNSHESAQSRRVVISHNMDKALKEVFDYSYRDYILSWYGNLSRDEGQLYHLLLEDFWEIA
RQLHHRLSHVDVVKVVCNDVVRTLLTHFCDLKAANARHEEQPRPFVLHACLRNSDDEVRFLQTCSRVLV
FCLLPSKDVQSLSLRIMLAEILTTKVLKPVVELLSNPDYINQMLLAQLAYREQMNEHHKRAYTYAPSYE
DFIKLINSNSDVEFLKQLRYQIVVEIIQATTISSFPQLKRHKGKETAAMKADLLRARNMKRYINQLTVA
KKQCEKRIRILGGPAYDQQEDGALDEGEGPQSQKMESHSVAQAGMQ

BR1-48

Nucleotide Sequence

GCCAGCCGGCCAATGTCTAAGCGAGGCGGAGCGGCCCAGGCGGCCCGAGCCTGGGGGAGCGCGCAGCCG
GCCAGTGGCGGCCTCGCCGGCGGCCTCTTCCCGGGCTCGCAGTAGGCCCGAGTCGTCGCCGGGAGCTCC
TGGGAGCAGCGTCCCCGCCCTGCTCCCCTCGCTCCCGCCTCTTGCGGCCCCACGGCCCCTCAGCGCCCG
CCCCCGGCTCCGCCCGCCGCAGCCGCAGCCCCTGGCGCTAACGGTCGGTAACGGCCCGCGCGCGCCGCC
CGCCGGGGGCTCGCGCCAGCCACGAGGGAGCGTCCGCGGCCCGCGCGCCCGCGCGGCGGAGGAGAGGTG
TTAAGTGTGATGCTTCCATAATACATTTGGATGCTGTCAGCTAAGTTCACTTCTGAACTAAGGGGTTCC
TCCAAATGTTGGCTGAAATTCATCCCAAGGCTGGTCTGCAAAGTCTGCAATTCATAATGGAGCTACTGT
ACTGGCTATTGGAAGGAGGAGATTCTGAAGATAAGGAGGTAAAACCTGTTTAGAAATTAAAAATGAGTT
ACGATTTAAAGAAAATTCAGATGACTCATTGTGAGTGCTAGTTCTCTTGTAGGATGCCACTGGAAATGT
TGAAATGAAAAATATTCAGCCGTTGGTCTTTGAAATTTCCTGTGATGTGTTTCAATCTAGATGCAAAGA
ACATGGAAAAATCAAAGTGCTCGAGTGGTTTAAATATGTTTTGGGTATTCCTGTTTATAGACTATAATA
CTTTTCCAATTAAAATCCTCAGTTGTCACGCAGAAGAAGGTTAAGCTGTATTTGATTGCCAGTTTTACT
GAAAATGCTTAGTATTTTACAGTATCACCAAATATATTTTGTTTAGCCAAGGTATAGGAAAAATAAAAT
AAATTGTATAGGTTGACTTTTTTCTAAAATGTCTTTATTGGATTGAATGAATGTTTATACCTGAAAAAA
AAAGGTTCAAAAAAAAAAAAAAAAGCTT

Predicted Protein Sequence

ASRPMSKRGGAAQAARAWGSAQPASGGLAGGLFPGSQ

BR1-49

Nucleotide Sequence

CAGAAATTAAGTATTGCAACTTACTGAGGGCCCTGGGAGATAGCATGGACGTGCATTGAGAAGCCAGCC
TCAGACCTTAGCTTGAAGCAGCTTGAGGCCAGACCTACTGTAGCCTCAGCATCTTGCTAGGAGGCATGG
AAGTGATCTATCCTGCCAGGAGGCCTCAGAGTGATCTGTCCTGCCAGGAGGTGTGAGAGTGATCTGTCC
TGTGAGGCATTTAGGGGCTTTCAGGAATTTAGTAAAAGGTGGAGTATGCCTTTCCAGTATCTTCCATCT
TCCTTTGTATACTTGTCCTTCCTCCCATTTCCTCCCTTTGGCCCGAGGTAGGAGGATGGAGGGAGGCTG
CTACTCTACCACTTCCTGTGTGCCTCTACTGTGGCCTCAACCCTGGCAATTATAGCTACTCCCATCCCT
TACCTGGGCATGTGTGAGCCCTTCTCACTGGATTTTATACCCTTGTGTCTGTGTACATAAATATATATA
CATATATATATACATAAAAACTTTGTACAAAAAAAAAAGCTT

Fig. 2B

BR1-52

Nucleotide Sequence 5'

GAATTCAAGCCTTCCGCGGACTCGCGCCGGGTGGCAGATGGCGGCGGTGCCGGGGGCACCTTCCAGCCC
TACCTAGACACCTTGCGGCAGGAGCTGCAGCAGACGGACCCAACGCTGTTGTCAGTAGTGGTGGCGGTT
CTTGCGGTGCTGCTGACGCTAGTCTTCTGGAAGTTAATCCGGAGCAGAAGGAGCAGTCAGAGAGCTGTT
CTTCTTGTTGGCCTTTGTGATTCCGGGAAAACGTTGCTCTTTGTCAGGTTGTTAACAGGCCTTTATAGA
GACACTCAGACGTCCATTACTGACAGCTGTGCTGTATACAGAGTCAACAATAACAGGGGCAATAGTCTG
ACCTTGATTGACCTTCCCGGCCATGAGAGTTTGAGGCTTCAGTTCTTAGAGCGGTTTAAGTCTTCAGCC
AGGGCTATTGTGTTTGTTGTGGATAGTGCAGCATTCCAGCGAGAGGTGAAAGATGTGGCTGAGTTTCTG
TATCAAGTCCTCATTGACAGTATGGGTCTGAAGAATACACCATCATTCTTAATAGCCTGCAATAAGCAA
GATATTGCAATGGCAAAATCAGCAAAGTTAATTCAACAGCAGCTGGAGAAAGAACTCAACACCTTACGA
GTTACCCGTTCTGCTGCCCCCAGCACACTGGACAGTTCCAGCACTGCCCCTGCTCAGCTGGGGAAGAAA
GGCAAAGAGTTTGAATTCTCACAGTTGCCCCTCAAAGTGGAGTTCCTGGAGTGCAGTGCCAAGGGTGGA
AGAGGGGACGTGGGCTCTGCTGACATCCAGGACTTGGAGAAATGGCTGGCTAAAATTGCCTGAGAGGCA
GCTCTAAAGCACAAGACCTGGATGTGTGACACACAGTTTTGGAAAAAGGTCTGTGGTAGTCTGGAGTTG
ATGAGGAAGGGGTACAAGATGTGGTTAGAAACATTTCTTTGTTCTGGAAACAAAGTACTGTTGAAACCA
GCTTGGAATTTTTTTTTTTTTTTTTTTTTAAGTTCAGTTCTCCCTTATGGCTGCCTTTCAAACAAGTACC
TTTTATCTGATGCCTGTATCTTCCCTTTGTTAAGGTGTAACTTGATGTAGGGTCAAGGTTTTTGTGACA
ACAGGCAGACTCCACACAGAGAGGATATGATGAGAATATGGCCATCACCTGAA

Predicted Protein Sequence

MASADSRRVADGGGAGGTFQPYLDTLRQELQQTDPTLLSVVVAVLAVLLTLVFWKLIRSRRSSQRAVLL
VGLCDSGKTLLFVRLLTGLYRDTQTSITDSCAVYRVNNNRGNSLTLIDLPGHESLRLQFLERFKSSARA
IVFVVDSAAFQREVKDVAEFLYQVLIDSMGLKNTPSFLIACNKQDIAMAKSAKLIQQQLEKELNTLRVT
RSAAPSTLDSSSTAPAQLGKKGKEFEFSQLPLKVEFLECSAKGGRGDVGSADIQDLEKWLAKIA

BR1-91

Nucleotide Sequence

GCCGCGCGTCTTGCCCTACCACCCTGCACAGCCCGGCCAGGCCGCCAAAAAGGCCGTCAGGACCCGCTA
CATCAGCACGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCC
CACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGACGGGCGC
ACGGGGCCGCCGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGACCCATTGGACACCT
GCACCTGGCGCTGCGCCACCTGCTGGAGCAGCACGGCGACGACTTTGACTGGTTCTTCCTGGTGCCTGA
CACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGGCCACCTCAGCCTGGCCTCCGCCGCCCA
CCTGTACCTGGGCCGGCCCCAGGACTTCATCGGCGGAGAGCCCACCCCCGGCCGCTACTGCCACGGAGG
CTTTGGGGTGCTGCTGTCGCGCATGCTGCTGCAACAACTGCGCCCCCACCTGGAAGGCTGCCGCAACGA
CATCGTCAGTGCGCGCCCTGACGAGTGGCT

Predicted Protein Sequence

PRVLPYHPAQPGQAAKKAVRTRYISTELGIRQRLLVAVLTSQTTLPTLGVAVNRTLGHRLERVVFLTGA
RGRRAPPGMAVVTLGEERPIGHLHLALRHLLEQHGDDFDWFFLVPDTTYTEAHGLARLTGHLSLASAAH
LYLGRPQDFIGGEPTPGRYCHGGFGVLLSRMLLQQLRPHLEGCRNDIVSARPDEW

Fig. 2C

BR1-95

Nucleotide Sequence

AGAAGGCCTGTCAGGTGAAGTCTTCAGCCTCCCACAGCGCAGGGTCCCAGCATCTCCACGCGCGCCCGT
GGGAGGTGGGTCCGGCCGGAGAGGCCTCCCGCGGACGCCGTCTCTCCAGAACTCCGCTTCCAAGAGGGA
GCCTTTGGCTGCTTTCTCTCCTTAAACTTAGATCAAATTTTTTGGTTTTTAATCAGTTATCTTGGGAAC
TTAACCTGGCCCCTCACCTCTTCTGCACCCCCCGCCCCCGAAACTGTCTCGTAATGAATTTCTGCTGTC
CTCCTGGGAGTGGACGGCCGGGTCCCGTCCCCCGGGAGCATCGCTCGGCTCAGCACCTTGGCTCCCAGT
GGGGGCCCCGTGGAGGGCGCCCGTAGTGATAAGCACACCGGCACGAACGTCAGGTCCATTCCTCGAAGT
CGGAGCCCTCACTCTGCCCTGTCCTGGGGCTGGCTGAGGGCGAACGCCCCACCTCACTTTCTAGAGCCC
TGTCTGTCCTAGCTCCTATCTGACCTTGTGTGTAAATACGTACATCTGTTTTTAAAGTGGATGGGCCCC
TGAGAACTCAGTGAAATGCAGAGTTCTCCATGCACC

BR1-102

Nucleotide Sequence

GGCGAGCGGGATGCTGCGGACTGCGCTGCGCGGCGCGCCGAGGTTGCTGAGTCGCGTGCAGCCCCGGGC
GCCCTGCCTGAGGCGGCTGTGGGGCCGCGGGGCCCGTCCAGAGGTCGCGGGGAGGCGGCGGGCCTGGGC
CTGGGGCTGGCGGCGCTCAAGCTCCGAGCAGGGGCCGGGGCCCGCGGCGGCTCTGGGGCGCGTGGAGGC
GGCGCACTACCAGCTCGTCTACACCTGCAAGGTCTGCGGGACTAGGTCCTCCAAGCGCATCTCCAAGCT
GGCCTATCACCAAGGCGTGGTCATTGTGACCTGCCCCGGCTGCCAGAACCACCATATCATCGCTGACAA
CCTGGGCTGGTTCTCGGACCTGAATGGGAAGAGAAATATCGAAGAGATCCTGACGGCCAGAGGCGAGCA
GGTGCACCGTGTGGCGGGCGAGGGGGCCCTGGAACTGGTTCTGGAGGCTGCAGGGGCCCCCACATCCAC
TGCAGCTCCGGAAGCGGGTGAGGATGAGGGTCCCCCCAGCCCTGGCAAGACGGAGCCGAGCTGACTCCT
GCGCTCCCCGGCACTGTGGGACTTCGGCCTTCCAGAAGGAGGTTTTGGCCCTGGGCCTCTCTGGACGTG
GCCTGTTTTCTATCAATAAACAGACATTACTTCCAAAAAAAAAAAAAAAAAAAGCTT

Predicted Protein Sequence

MLRTALRGAPRLLSRVQPRAPCLRRLWGRGARPEVAGRRRAWAWGWRRSSSEQGPGPAAALGRVEAAHY
QLVYTCKVCGTRSSKRISKLAYHQGVVIVTCPGCQNHHIIADNLGWFSDLNGKRNIEEILTARGEQVHR
VAGEGALELVLEAAGAPTSTAAPEAGEDEGPPSPGKTEPS

BR1-105

Nucleotide Sequence

GGAGGCCAGCCCCGGACCCCCATTCGCGCCCAGGGCCCGCAGGAGGCTGGGCAGCGGCCCGGACAGGGA
GCTCCGCAAGCCGGAGGAGCCGGAGAACGGCGAGCCCACGGCTGCGGCCACCGCCAGGAGGAGCAAGAG
GGAGAGGCGCGAGGAGGACAGGGCCCCGGCAGAGCAGGTCCCGCGGAGCCCGGTCATCAAGATCTCCTA
CAGCACGCCCCAGGGCAAGGGAGAGGTGGTCAAGATCCCCTCCCGCGTGCACGGCTCTCTGGAGCCCTT
CCGTCCCCAGCAGGCCCCGCAGGACGACGGCAGCCAGGACCCCGAGGTGCTGGACAGAGAGTCCCGGGA
CCGGCCGTCCTGCGCGCCCTCGGCCTCCATCCCCAAGTTGAAACTGACACGGCCTGTGCCGGCCGGCGC
GGACCTGCCGCCCCCTAAGATCCGCCTGAAGCCCCACCGTCTGGGGGACAGCGAGCACGAGCCCGTGTA
CCGGGCCGAGCTGGTGGGGGAGCTGAACGGGTACCTGCGGGACAGCTCGCCGGCGCCCTGTGCGGACGG
CCCTGCCGGTGGGCTGGCGGACTTGTCTTCTGG

Fig. 2D

Predicted Protein Sequence

EASPGPPFAPRARRRLGSGPDRELRKPEEPENGEPTAAATARRSKRERREEDRAPAEQVPRSPVIKISY
STPQGKGEVVKIPSRVHGSLEPFRPQQAPQDDGSQDPEVLDRESRDRPSCAPSASIPKLKLTRPVPAGA
DLPPPKIRLKPHRLGDSEHEPVYRAELVGELNGYLRDSSPAPCADGPAGGLADLSS

BR1-109

Nucleotide Sequence

GAATCAAAGCACAATCAAGAACTGACATCTCAGTTGTTAGCTGCAGAAAATAAATGCAATCTATTAGAA
AAACAATTGGAATACATGCGAAATATGATAAAGCATGCCGAAATGGAGAGGACATCTGTCTTAGAGAAA
CAAGTTTCCCTAGAAAGAGAACGACAACATGATCAAACACATGTTCAGAGCCAACTTGAAAAATTGGAT
CTTCTTGAACAGGAGTATAACAAACTTACCACAATGCAGGCCCTTGCAGAAAAAAAAATGCAAGAGTTG
GAAGCAAAACTCCATGAAGAAGAACAGGAAAGGAAACGCATGCAAGCTAAGGCAGCTGAGGTAAGTTAA
AATGTGAGAAAGTGGGCTCTTCATATTTCTTACCGCTTTTTGTGTACAAGTAAACAATTACATTTAGGT
ATCTTACATTATTTGTATTTTAGAATAGTCCAACATACAAATTTTCAGATATAGGTTAATGTAGAAAAT
TATGAACAAGTATTTTTATAGACAGAGTAAAATTTGAACAGTTTTTTAGCATGTTTGTAATTGTAGAGT
TTCAATCAGCATTTTAAAAAATAACAGA

Predicted Protein Sequence

ESKHNQELTSQLLAAENKCNLLEKQLEYMRNMIKHAEMERTSVLEKQVSLERERQHDQTHVQSQLEKLD
LLEQEYNKLTTMQALAEKKMQELEAKLHEEEQERKRMQAKAAEVS

BR1-111

Nucleotide Sequence

CCAAGACGGACCAAGCAGATGGACCCAGAGAGCCACCGCAGAGTGCCAGGAGGAAGCGCAGCTACAAGC
AGGCCGTGAGTGAGCTGGACGAGGAGCAGCACCTGGAGGATGAGGAGCTGCAGCCCCCCAGGAGCAAGA
CCCCTTCCTCACCCTGCCCAGCCAGCAAGGTGGTACGGCCCCTCCGGACCTTTCTGCACACAGTGCAGA
GGAACCAGATGCTCATGACCCCGACCTCAGCCCCACGCAGCGTCATGAAGTCCTTTATTAAGCGCAACA
CTCCCCTGCGCATGGACCCCAAGGAGAAGGAGCGGCAGCGCCTGGAGAATCTGCGGCGGAAGGAGGAGG
CCGAGCAGCTGCGCAGGCAGAAGGTGGAGGAGGACAAGCGGCGGCGGCTGGAGGAGGTGAAGCTGAAGC
GTGAGGAACGCCTCCGCAAGGTGCTGCAGGCCCGCGAGCGGGTGGAGCAGATGAAAGAAAGAAAAAAAA
AAAAAGCTT

Predicted Protein Sequence

KTDQADGPREPPQSARRKRSYKQAVSELDEEQHLEDEELQPPRSKTPSSPCPASKVVRPLRTFLHTVQR
NQMLMTPTSAPRSVMKSFIKRNTPLRMDPKEKERQRLENLRRKEEAEQLRRQKVEEDKRRRLEEVKLKR
EERLRKVLQARERVEQM

Fig. 2E

BR1-188

Nucleotide Sequence

CGTTTTGGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAAAAAAAAACAAGT
ATAATGGAAAAAAAAAAAAAAGAATATGAGGGTTGTCCTTTCTGCCTTGGATGAACTATGGTCATCCTGA
GCCTTTTTTTTTTTTTTTTGCAGCCCCAAAGGGGCAAAAAGAGACTTTAATTAGGGGAGGGAGGATCCA
CCAGAATCAGAAAAGGGACAGCTAGCGTGGGAGCAGAGGAGCCAGAACAGGCAGGAGGAGGGCCCGGCC
AGGAAGCTCTGGAGGACTCACCTCGCCACCTCTGGCACAGGCACTGGCACTGACGGACAAGGCGAAACA
GCGGCCCCTCTCAACTGGGAGGGCACCCAATGGCCCCTGTAGCCAGAGGTTGCCCGGCTTTTGGGCCCC
AGGTCCTAGGCATGACTGGTGGTCACCAATTTGGCCCTTGTCCCCAACCAGTGCTGGGGGGCCATCTTT
AGGCAGAACTCAGGAAGCTGGCAGAGAGTTCCCAACACTTATACTGGGGCCAAGGAGGCTTGGGAATAA
AAGGGCAAGGGAGAAACTGGAGTCTGAAACCTCCCTCCCTCCCCACCTGGGGTGGCATAGGAGATAAAA
AGAGGGGAGTATAAAAAAAAAAAAAAAAGCTT

Predicted Protein Sequence

FWYSIGNVNELNT

Fig. 2F

BR1-74

Nucleotide Sequence

TGGCCCTCGAGGCCAAGAATTCGGCACGAGGCGCGTCTCCAGTCCTCGCACCTGGAACCCCAACGTCCC
CGAGAGTCCCCGAATCCCCGCTCCCAGGCTACCTAAGAGGATGAGCGGTGCTCCGACGGCAGGGGCAGC
CCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGCTCAGGGCGGACCCGTGCAGTCCAAGTCGCC
GCGCTTTGCGTCCTGGGACGAGATGAATGTCCTGGCGCACGGACTCCTGCAGCTCGGCCAGGGGCTGCG
CGAACACGCGGAGCGCACCCGCAGTCAGCTGAGCGCGCTGGAGCGGCGCCTGAGCGCGTGCGGGTCCGC
CTGTCAGGGAACCGAGGGGTCCACCGACCTCCCGTTAGCCCCTGAGAGCCGGGTGGACCCTGAGGTCCT
TCACAGCCTGCAGACACAACTCAAGGCTCAGAACAGCAGGATCCAGCAACTCTTCCACAAGGTGGCCCA
GCAGCAGCGGCACCTGGAGAAGCAGCACCTGCGAATTCAGCATCTGCAAAGCCAGTTTGGCCTCCTGGA
CCACAAGCACCTAGACCATGAGGTGGCCAAGCCTGCCCGAAGAAAGAGGCTGCCCGAGATGGCCCAGCC
AGTTGACCCGGCTCACAATGTCAGCCGCCTGCACCGGCTGCCCAGGGATTGCCAGGAGCTGTTCCAGGT
TGGGGAGAGGCAGAGTGGACTATTTGAAATCCAGCCTCAGGGGTCTCCGCCATTTTTGGTGAACTGCAA
GATGACCTCAGATGGAGGCTGGACAGTAATTCAGAGGCGCCACGATGGCTCAGTGGACTTCAACCGGCC
CTGGGAAGCCTACAAGGCGGGGTTTGGGGATCCCCACGGCGAGTTCTGGCTGGGTCTGGAGAAGGTGCA
TAGCATCATGGGGGACCGCAACAGCCGCCTGGCCGTGCAGCTGCGGGACTGGGATGGCAACGCCGAGTT
GCTGCAGTTCTCCGTGCACCTGGGTGGCGAGGACACGGCCTATAGCCTGCAGCTCACTGCACCCGTGGC
CGGCCAGCTGGGCGCCACCACCGTCCCACCCAGCGGCCTCTCCGTACCCTTCTCCACTTGGGACCAGGA
TCACGACCTCCGCAGGGACAAGAACTGCGCCAAGAGCCTCTCTGGAGGCTGGTGGTTTGGCACCTGCAG
CCATTCCAACCTCAACGGCCAGTACTTCCGCTCCATCCCACAGCAGCGGCAGAAGCTTAAGAAGGGAAT
CTTCTGGAAGACCTGGCGGGGCCGCTACTACCCGCTGCAGGCCACCACCATGTTGATCCAGCCCATGGC
AGCAGAGGCAGCCTCCTAGCGTCCTGGCTGGGCCTGGTCCCAGGCCCACGAAAGACGGTGACTCTTGGC
TCTGCCCGAGGATGTGGCCGTTCCCTGCCTGGGCAGGGGCTCCAAGGAGGGGCCATCTGGAAACTTGTG
GACAGAGAAGAAGACCACGACTGGAGAAGCCCCCTTTCTGAGTGCAGGGGGGCTGCATGCGTTGCCTCC
TGAGATCGAGGCTGCAGGATATGCTCAGACTCTAGAGGCGTGGACCAAGGGGCATGGAGCTTCACTCCT
TGCTGGCCAGGGAGTTGGGGACTCAGAGGGACCACTTGGGGCCAGCCAGACTGGCCTCAATGGCGGACT
CAGTCACATTGACTGACGGGGACCAGGGCTTGTGTGGGTCGAGAGCGCCCTCATGGTGCTGGTGCTGTT
GTGTGTAGGTCCCCTGGGGACACAAGCAGGCGCCAATGGTATCTGGGCGGAGCTCACAGAGTTCTTGGA
ATAAAAGCAACCTCAGAACAAAAAAAAAAGCTT

Predicted Protein Sequence

MSGAPTAGAALMLCAATAVLLSAQGGPVQSKSPRFASWDEMNVLAHGLLQLGQGLREHAERTRSQLSAL
ERRLSACGSACQGTEGSTDLPLAPESRVDPEVLHSLQTQLKAQNSRIQQLFHKVAQQQRHLEKQHLRIQ
HLQSQFGLLDHKHLDHEVAKPARRKRLPEMAQPVDPAHNVSRLHRLPRDCQELFQVGERQSGLFEIQPQ
GSPPFLVNCKMTSDGGWTVIQRRHDGSVDFNRPWEAYKAGFGDPHGEFWLGLEKVHSIMGDRNSRLAVQ
LRDWDGNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPPSGLSVPFSTWDQDHDLRRDKNCAKSL
SGGWWFGTCSHSNLNGQYFRSIPQQRQKLKKGIFWKTWRGRYYPLQATTMLIQPMAAEAAS

Fig. 2G

BR1-72

Nucleotide Sequence

CTTTCTTTTAATTGGAGTGAGGACAAAACCTGATCCCGTTCAGAGATGAACAGAATCATCTAGGTCGTT
AAGAACAGTCTTATCTTTCTGGTTAACCAGGTCCTAAATTCAAGGGATTTCCACGTGTTCTTTAAAACC
ACCTGAAACTGTCCTGCGCAGGCTCACACTGCGGCTCTGTACCAAGGGCTTGGCCCCGCCTGCAGCCAG
AGGGCATCTCACTGGATGTCATCATCATCAAAAAGATCTTCAAAATCGTTAAAAAAGTCTGCATGAACG
GCCTTTTCATTGGCTGCCAAGTCCATCAAGAGCCCGGTGCTGCCTCCCGCCTCGTCCAGGTCCACGTAG
GGCCCGGCGCCCTCGGGGAACATCTCGTCCACCGCCGGCTTCATCTCGGGGCCGCGGCGCTCTAGGCTC
ACTTCCGGCCTCAGAGCCGCTTCCGGCGCGCGCCACATGGCAGATCTAGACGCCGGAGGACCGCGCCTC
TGTGGTGCCCGCCTGCTGGGGTGCTCCCGGCCGCCGGCGCGTCCGCGGCCACCGGCGTTTGACTTTGCG
CTGGCGAGGGCGGGCGTCTGCAACGCCTCTGATGG

BR1-77

Nucleotide Sequence 5'

GGGACCCGTCCGACCTCAAGGAGGCGGTCACATACATCCGCTTCCGACACCCGGCGGCGCCGCTGTTCG
CGGTGAGCGAAGGCTCGGGCTCGGCGCTGCTCCTGTCCTACCTGGGCGAGTGCGGCTCCTCCAGCTACG
TGACAGGCGCCGCCTGCATCTCGCCCGTGCTGCGCTGCCGAGAGTGGTTCGAGGCCGGCCTGCCCTGGC
CCTACGAGCGGGGCTTTCTGCTCCACCAGAAGATCGCCCTCAGCAGGTATGCCACAGCCCTGGAGGACA
CTGTGGACACCAGCAGACTGTTCAGGAGCCGTTCCCTTCGAGAGTTTGAGGAGGCTCTCTTCTGCCACA
CCAAAAGCTTCCCCATCAGCTGGGATACCTACTGGGACCGCAACGACCCGCTCCGGGATGTCGATGAGG
CAGCCGTGCCTGTGCTGTGTATCTGCAGTGCTGACGACCCCGTGTGTGGACCCCCAGACCACACTCTGA
CAACTGAACTCTTCCACAGCAACCCCTACTTCTTCCTCCTGCTCAGTCGCCACGGAGGCCACTGTGGCT
TCCTGCGCCAGGAGCCCTTGCCAGCCTGGAGCCATGAGGTCATCTTGGAGTCCTTCCGGGCCTTGACTG
AGTTCT

BR1-82

Nucleotide Sequence

AGAAGGCCTGTCAGGTGAAGTCTTCAGCCTCCCACAGCGCAGGGTCCCAGCATCTCCACGCGCGCCCGT
GGGAGGTGGGTCCGGCCGGAGAGGCCTCCCGCGGACGCCGTCTCTCCAGAACTCCGCTTCCAAGAGGGA
GCCTTTGGCTGCTTTCTCTCCTTAAACTTAGATCAAATTTTTTGGTTTTTAATCAGTTATCTTGGGAAC
TTAACCTGGCCCCTCACCTCTTCTGCACCCCCCGCCCCCGAAACTGTCTCGTAATGAATTTCTGCTGTC
CTCCTGGGAGTGGACGGCCGGGTCCCGTCCCCCGGGAGCATCGCTCGGCTCAGCACCTTGGCTCCCAGT
GGGGGCCCCGTGGAGGGCGCCCGTAGTGATAAGCACACCGGCACGAACATCAGGTCCATTCCTCGAAGT
CGGAGCCCTCACTCTGCCCTGTCCTGGGGCTGGCTGAGGGCGAACGCCCCACCTCACTTTCTAGAGCCC
TGTCTGTCCTAGCTCCTATCTGACCTTGTGTGTAAATACGTACATCTGTTTTTAAAGTGGATGGGCCCC
TGAGAACTCAGTGAAATGCAGAGTTCTCCATGCACCTAAAGCTCCTTTGTCGCTCTCATGGCTGTCAGA
TCCTGGTCCCTCCACACTGGGTGCTGGGGAGGGAGGACCCTCGGGGCTACCGCGCGCCCCCCCATCCCA
CAGATCAGGAGCCAAGGAGGGAGAACAGGGCAGCCTGTGGGACTCTAGGATGCTTCAGAAGAAGCGACG
GCACCGTCAACCCTCTGTTTTTTAAAGGTGGTTGGAGACTGTTAACACTGAGCTCATTGACTTCTAAAA
AAAAAAAGCTT

Fig. 2H

BR1-85

Nucleotide Sequence 5'

GGGACACGCGCGAGGAGCCGCCGCAGCCGCCGCCGCCGCGCTGTGGAGCCCGAGTGAGCGCGGCGCCAG
CCCGTCCGGCCGCTGGACAACATGGAGGCTGCACTGCCCGGGCCGCCGTGGCCGCTGCTGCTGGCGCTG
TGCGGCTGCCCGGCCCCCGCCGCAGGTACATGTACTGGGCAGATTGGGGTGAGACGCCCTGGATCGAGC
AGGTGGGGATGGACGACAGCACCCAGAAGATCATTGTGGACTTGGACATTTACTGGCCCAATAGACTGA
CCATCGACCTGGGGGAGAAGAAGCTCTACTGGGCTGACACTAAGCTCGGCTTCATCCACCGTGCCAACC
TGGACGGCTCGTTCCGGCAGAAGGTGGTGAAGGGCAGCCTGACGCACCCCTTCACCCTGACACTCTCCG
GGGACACTCTGCACTGGATGGACTGGCTGACGTGCTCCATCCATGCCTGCAACAGGAACACTGCGGGGA
AGAGGAACGAGATCTTGAGCGCCCTCTACTCAGCATGGATATCCAGGTGCTGAGCCAGGAGCAGCAGCC
TTTCTGAGCCAAGGAGGTCCTGCTGCTGGCCCG

BR1-90

Nucleotide Sequence 5'

TGGAGCAGCCCCACCACAAGAAGGAGTGCTACCTGAACTTCGATGACACAGTGTTCTGCGACAGCGTAT
TGGCCACCAACGTGACCCAGCAGGAGTGCTGCTGCTCTCTGGGGGCCGGCTGGGGCGACCACTGCGAAA
TCTACCCCTGCCCAGTCTACAGCTCAGCCGAGTTCCACAGCCTCTGCCCAGACGGAAAGGGCTACACCC
AGGACAACAACATCGTCAACTACGGCATCCCAGCCCACCGTGACATCGACGAGTGCATGTTGTTCGGGT
CGGAGATTTGCAAGGAGGGCAAGTGCGTGAACACGCAGCCTGGCTACGAGTGCTACTGCAAGCAGGGCT
TCTACTACGACGGGAACCTGCTGGAATGCGTGGACGTGGACGAGTGCCTGGACGAGTCCAACTGCCGGA
ACGGAGTGTGTGAGAACACGCGCGGCGGCTACCGCTGTGCCTGCACGCCCCCTGCCGAGTACAGTCCCG
CGCAGCGCCAGTGCCTGAGCCCGGAAGAGATGGACGTGGACGAGTGCCAGGACCCGGCAGCCTGCCGCC
CTGGCCGCTGCGTCAACCTGCCGGGCTCCTAC

Predicted Protein Sequence

EQPHHKKECYLNFDDTVFCDSVLATNVTQQECCCSLGAGWGDHCEIYPCPVYSSAEFHSLCPDGKGYTQ
DNNIVNYGIPAHRDIDECMLFGSEICKEGKCVNTQPGYECYCKQGFYYDGNLLECVDVDECLDESNCRN
GVCENTRGGYRCACTPPAEYSPAQRQCLSPE

BR1-204

Nucleotide Sequence 5'

CTCGTCAGGAGCCAACGTAAGAGTTCTTCACTACCACAACCAAGGAGGGATATGATAGGCGGCCAGTGG
ATATAACTCCTTTAGAACAAAGGAAATTAACTTTTGATACCCATGCATTGGTTCAGGACTTGGAAACTC
ATGGATTTGACAAAACACAAGCAGAAACAATTGTATCAGCGTTAACTGCTTTATCAAATGTCAGCCTGG
ATACTATCTATAAAGAGATGGTCACTCAAGCTCAACAGGAAATAACAGTACAACAGCTAATGGCTCATT
TGGATGCTATCAGGAAAGACATGGTCATCCTAGAGAAAAGTGAATTTGCAAATCTGAGAGCAGAGAATG
AGAAAATGAAAATTGAATTAGACCAAGTTAAGCAACAACTAATGCATGAAACCAGTCGAATCAGAGCAG
ATAATAAACTGGATATCAACTTAGAAAGGAGCAGAGTAACAGATATGTTTACAGATCAAGAAAAGCAAC
TTATGGA

Fig. 2I

Predicted Protein Sequence

EFFTTTTKEGYDRRPVDITPLEQRKLTFDTHALVQDLETHGFDKTQAETIVSALTALSNVSLDTIYKEM
VTQAQQEITVQQLMAHLDAIRKDMVILEKSEFANLRAENEKMKIELDQVKQQLMHETSRIRADNKLDIN
LERSRVTDMFTDQEKQLM

BR1-207

Nucleotide Sequence

CAGACGTCTGGGATTCCCCAGAAGGCTCTGACACCCTCTGCCCGCCCTGTAGCTGTAGTCCTCCCATTG
GCTAGGGCTCTGGGGTCGGGCAGGTTTCGGGTGCCCCCAGTGGGCCTCTGGTTCCAGGCAGCTCGTGAC
AAGCCCCTGTGCTCTCTAGAAAGCCCGTTTTGGCCTGAGTGCGGCTGAGGACATCACCCCCCGGTTCAG
GGCAGCCTGTGAGCAGCAAACTGTGGCTCTGACTCTGCAGGAGGACAGAGCATCCCTGACGCTTTCAGG
GGGGCCCTCGGCACTGGCCTTTGACCTCTCCAAGGTACCAGGCCCAGAGGCAGCCCCCAGGCTGCGGGC
GCTGACACTGGGCCTGGCAAAACGCGTGTGGAGCCTGGAGCGGCGACTGGCAGCTGCAGAAGAGACAGC
TGTCAGCCCGAGGAAGAGCCCCCGGCCTGCAGGGCCTCAGCTCTTCTTACCAGACCCAGATCCCCAGAG
AGGTGGCCCTGGACCTGGAGTCAGGAGGCGGTGTCCAGGAGAGTCGCTCCTCAACCCCGGGTTCAAGAG
TAAGAAACCAGCTGGTGGCGTGGACTTCGATGAGACCTGAAGGTGCAGCACAAGCGTGGCCCCGCGGGG
AGTCCGCCTATGAGGGGAGAGGCAGTCTTTGAGGCCCCCATCAGAGACCCCCCGCCACCACCTCCACCT
GCCTGTCCTGGGCCAGGACTAACACGGCTCCTCAAATTCCTTCCCTGTCAAATAAACAGCTCCCTTGGT
TGGAAAAAAAAAAAAAAAAAAGCTT

BR1-214

Nucleotide Sequence 5'

AAAAAAAAAGGAAAAGTACCCAGTGCTCTCAGCTTCTGAGCCTCCTCTACAGCCCTGTTCACTTTTAAA
CCTGTGCCCTGTGTCTGTGTCCCCACTTAATATATATAGTACACAGCTGGAGAGATGGCTCAGCCAGGA
GAGGGACCCATAGGTCTGTGAATTCCAGAGGAGAGGCAGGTATTTAGAGGTGGCTCTGTCAGGTGAAAT
CTGAGGAGCCAAAGCTATGTATGTGCATATGTCAGCCGGGCTCTGTGGGAGGTGGTGTAGACCTATGGC
ATGGGACAGGTGTGCACGCTGGGATCTCTGGCCGGTTCCGAAAAGTGAGGATCAGGTAGTGGGTGGCTG
ATTGCACAAGTTTAGAACCCAGGATTAGGGACACACAGGTCAGCACCTGCTTCTCAGCATCCTGACTGG
GTGTGATGGGCATACTCAAGGC

BR1-215

Nucleotide Sequence 3'

GACCCAGTTTCATGGAAGAGAATTTGTCCATGGATGGAGTGGGGATGGGGGGTGGTTTGGGGATGAAAA
CTGTTCCATCTCAGATGAACAGGCATTAGATTCTCAATAAGGAGTGCGCAACCTAGATCCCTTGCATGT
ACAGTTCACAACAGGGTTCGCGCTCCTATGAGAATCTAATGCTGCTGCTGATCTGACAGGGGGCGAAGC
TCAGGTGGTAATGCTTGGCTGTCTAGGTCTGCTGCCCCATGATTGGGGACCCCCCTGCTCTACATCACT
AAAACCTTTACCATCTGGCCCAAATCCACCTCTCCTCCTTCATGTTTTACCAATTCCACCTACTCATGG
CTTCCAGGAAGCATTTTCTATAATCTACCTCCAGCCTAGCTACTTAAGTGACCATTCCCATTATATCTT
ATAATGTCTGTCCCTGCACCTACCAAATTGTATTTATTTATTTATTTATTTATTTAATTTATTTATTTT
TAGAGACAGTCTCTCCTCAGCCTGCCGAGTAGCTAGGACTACAGGCACACACCACCATGCCCAGCTAAT
TAAAAGCTT

Fig. 2J

SEQUENCE LISTING

<110> Corixa Corporation
Lodes, Michael J.

<120> COMPOSITIONS AND METHODS FOR BREAST CANCER THERAPY AND DIAGNOSIS

<130> 210121.473PC

<140> PCT
<141> 2000-02-29

<160> 192

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 38
<212> DNA
<213> Homo sapien

<400> 1
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 2
<211> 38
<212> DNA
<213> Homo sapien

<400> 2
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 3
<211> 38
<212> DNA
<213> Homo sapien

<400> 3
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 4
<211> 38
<212> DNA
<213> Homo sapien

<400> 4
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 5
<211> 38
<212> DNA
<213> Homo sapien

<400> 5
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 6
<211> 38
<212> DNA
<213> Homo sapien

<400> 6
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 7
<211> 38
<212> DNA
<213> Homo sapien

<400> 7
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 8
<211> 38
<212> DNA
<213> Homo sapien

<400> 8
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 9
<211> 38
<212> DNA
<213> Homo sapien

<400> 9
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 10
<211> 38
<212> DNA
<213> Homo sapien

<400> 10
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 11
<211> 38
<212> DNA
<213> Homo sapien

<400> 11
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 12
<211> 38
<212> DNA

<213> Homo sapien

<400> 12
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 13
<211> 38
<212> DNA
<213> Homo sapien

<400> 13
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 14
<211> 38
<212> DNA
<213> Homo sapien

<400> 14
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 15
<211> 38
<212> DNA
<213> Homo sapien

<400> 15
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 16
<211> 38
<212> DNA
<213> Homo sapien

<400> 16
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 17
<211> 38
<212> DNA
<213> Homo sapien

<400> 17
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 18
<211> 38
<212> DNA
<213> Homo sapien

<400> 18
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 19
<211> 38

<212> DNA
<213> Homo sapien

<400> 19
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 20
<211> 38
<212> DNA
<213> Homo sapien

<400> 20
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 21
<211> 38
<212> DNA
<213> Homo sapien

<400> 21
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 22
<211> 38
<212> DNA
<213> Homo sapien

<400> 22
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 23
<211> 38
<212> DNA
<213> Homo sapien

<400> 23
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 24
<211> 38
<212> DNA
<213> Homo sapien

<400> 24
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 25
<211> 38
<212> DNA
<213> Homo sapien

<400> 25
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 26

<211> 38
<212> DNA
<213> Homo sapien

<400> 26
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 27
<211> 38
<212> DNA
<213> Homo sapien

<400> 27
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 28
<211> 38
<212> DNA
<213> Homo sapien

<400> 28
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 29
<211> 38
<212> DNA
<213> Homo sapien

<400> 29
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 30
<211> 38
<212> DNA
<213> Homo sapien

<400> 30
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 31
<211> 38
<212> DNA
<213> Homo sapien

<400> 31
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 32
<211> 38
<212> DNA
<213> Homo sapien

<400> 32
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 33
<211> 38
<212> DNA
<213> Homo sapien

<400> 33
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 34
<211> 38
<212> DNA
<213> Homo sapien

<400> 34
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 35
<211> 38
<212> DNA
<213> Homo sapien

<400> 35
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 36
<211> 38
<212> DNA
<213> Homo sapien

<400> 36
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 37
<211> 38
<212> DNA
<213> Homo sapien

<400> 37
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 38
<211> 38
<212> DNA
<213> Homo sapien

<400> 38
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 39
<211> 38
<212> DNA
<213> Homo sapien

<400> 39
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 40
<211> 38
<212> DNA
<213> Homo sapien

<400> 40
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 41
<211> 38
<212> DNA
<213> Homo sapien

<400> 41
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 42
<211> 38
<212> DNA
<213> Homo sapien

<400> 42
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 43
<211> 38
<212> DNA
<213> Homo sapien

<400> 43
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 44
<211> 38
<212> DNA
<213> Homo sapien

<400> 44
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 45
<211> 38
<212> DNA
<213> Homo sapien

<400> 45
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 46
<211> 38
<212> DNA
<213> Homo sapien

<400> 46

cagtagctag catgcggacg actactacta cgacgacg 38

<210> 47
<211> 38
<212> DNA
<213> Homo sapien

<400> 47
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 48
<211> 38
<212> DNA
<213> Homo sapien

<400> 48
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 49
<211> 38
<212> DNA
<213> Homo sapien

<400> 49
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 50
<211> 38
<212> DNA
<213> Homo sapien

<400> 50
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 51
<211> 38
<212> DNA
<213> Homo sapien

<400> 51
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 52
<211> 38
<212> DNA
<213> Homo sapien

<400> 52
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 53
<211> 38
<212> DNA
<213> Homo sapien

<400> 53
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 54
<211> 38
<212> DNA
<213> Homo sapien

<400> 54
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 55
<211> 38
<212> DNA
<213> Homo sapien

<400> 55
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 56
<211> 38
<212> DNA
<213> Homo sapien

<400> 56
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 57
<211> 38
<212> DNA
<213> Homo sapien

<400> 57
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 58
<211> 38
<212> DNA
<213> Homo sapien

<400> 58
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 59
<211> 38
<212> DNA
<213> Homo sapien

<400> 59
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 60
<211> 38
<212> DNA
<213> Homo sapien

<400> 60
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 61
<211> 38
<212> DNA
<213> Homo sapien

<400> 61
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 62
<211> 38
<212> DNA
<213> Homo sapien

<400> 62
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 63
<211> 38
<212> DNA
<213> Homo sapien

<400> 63
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 64
<211> 38
<212> DNA
<213> Homo sapien

<400> 64
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 65
<211> 38
<212> DNA
<213> Homo sapien

<400> 65
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 66
<211> 38
<212> DNA
<213> Homo sapien

<400> 66
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 67
<211> 38
<212> DNA

<213> Homo sapien

<400> 67
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 68
<211> 38
<212> DNA
<213> Homo sapien

<400> 68
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 69
<211> 38
<212> DNA
<213> Homo sapien

<400> 69
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 70
<211> 38
<212> DNA
<213> Homo sapien

<400> 70
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 71
<211> 38
<212> DNA
<213> Homo sapien

<400> 71
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 72
<211> 38
<212> DNA
<213> Homo sapien

<400> 72
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 73
<211> 38
<212> DNA
<213> Homo sapien

<400> 73
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 74
<211> 38

<212> DNA
<213> Homo sapien

<400> 74
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 75
<211> 38
<212> DNA
<213> Homo sapien

<400> 75
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 76
<211> 38
<212> DNA
<213> Homo sapien

<400> 76
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 77
<211> 38
<212> DNA
<213> Homo sapien

<400> 77
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 78
<211> 38
<212> DNA
<213> Homo sapien

<400> 78
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 79
<211> 38
<212> DNA
<213> Homo sapien

<400> 79
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 80
<211> 38
<212> DNA
<213> Homo sapien

<400> 80
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 81

<211> 38
<212> DNA
<213> Homo sapien

<400> 81
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 82
<211> 38
<212> DNA
<213> Homo sapien

<400> 82
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 83
<211> 38
<212> DNA
<213> Homo sapien

<400> 83
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 84
<211> 38
<212> DNA
<213> Homo sapien

<400> 84
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 85
<211> 38
<212> DNA
<213> Homo sapien

<400> 85
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 86
<211> 38
<212> DNA
<213> Homo sapien

<400> 86
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 87
<211> 38
<212> DNA
<213> Homo sapien

<400> 87
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 88
<211> 38
<212> DNA
<213> Homo sapien

<400> 88
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 89
<211> 38
<212> DNA
<213> Homo sapien

<400> 89
cagtagctag catgcggacg actactacta cgacgacg 38

<210> 90
<211> 441
<212> DNA
<213> Homo sapien

<400> 90
cgtgcacaat tcattaaatg aacgtctcaa aaattaaaaa aattataaga tacgtatttc 60
tttaggcctt tgtgttttta aattaaaacc aacaaaaaga agtctccctc tccactccac 120
ccagcagcaa gggcagccgg aacgcttcgc tccagctacc tggcctcccg caagagggtt 180
cccccatgag accgttagtc tctctttgcc tggctgacta cctgcataca gtaggcactc 240
actgctggag tgaggcactg actcctccaa agattgcagg gggcggagga gggaaccacg 300
aaggcctggg agggggcatc tttggccccc actaaccatc tccctatttc tgcatcctgg 360
tgaccgtcag caagagatga gtcggggaga ccctctcctg gagttctagc ccctaattct 420
gggctttcta tatgagagga c 441

<210> 91
<211> 441
<212> DNA
<213> Homo sapien

<400> 91
cccaacatga acaagttgaa atcatcgcag aaggataaag ttcgtcagtt tatgatcttc 60
acacaatcta gtgaaaaaac agcagtaagt tgtctttctc aaaatgactg gaagttagat 120
gttgcaacag ataatttttt ccaaaatcct gaactttata tacgagagag tgtaaaagga 180
tcattggaca ggaagaagtt agaacagctg tacaatagat acaaagaccc tcaagatgag 240
aataaaattg gaatagatgg catacagcag ttctgtgatg acctggcact cgatccagcc 300
agcattagtg tgttgattat tgcgtggaag ttcagagcag caacacagtg cgagttctcc 360
aaacaggagt tcatggatgg catgacagaa ttaggatgtg acagcataga aaaactaaag 420
gcccagatac ccaagatgga a 441

<210> 92
<211> 431
<212> DNA
<213> Homo sapien

<400> 92
ccgaattcaa gcgagctatc tgctgggctg tgggaaagag aaatgcagtc tcctgcacgg 60
gagggtgtga gaaggccacc caggcagcct gcagctagag ctcggcgctg agagagcaca 120
ggagagtcag acacctccat tctggttaca ctttttccag atgctgaatg gagactcgct 180

gggaaagccc gcctccagac aatttcaaca tagcgccggt aacaccctag ttttgctcaa 240
actctgactt gtgctatctg tcccagacag ttcacgactg cccgtcactt ccctccttca 300
cactgcacca ggaaaaggcc acacctccct acacagcagc agccttttag gaaatgtgct 360
ctacaggaaa agtggttttt ctgattaaac atcaggaact agagtcaaca atcgctgtcc 420
tctcccgtgc c 431

<210> 93
<211> 441
<212> DNA
<213> Homo sapien

<400> 93
cgttgccacg aagttacttc taaaacaagg ccagtccccc acgtgaaagg tggccggtgc 60
agttgtgagg ctcgcttatt tgcctcgaag cataggtcac tgtcaggtcc taagggatca 120
ggcctagggg atcagccctc cttcaggtca tggtccagtg gtcctggagg ccacctctaa 180
ggccagcagt gctccttggg ccattaccag tgggagccag gctaagctgc ctgaccaaat 240
gttgacctac cctgccttat gccaggcagg actgcacgtg agcattgcca tgtggcatgg 300
caactgattc ggaagaggca aaacagcctt ctctccagga tgggcatgca gaggcctaaa 360
aaaaagccga agttcttttc cagggctaac gagaataaga ttagacccct ggaaaagtcc 420
tactaatggt taccatgaca g 441

<210> 94
<211> 441
<212> DNA
<213> Homo sapien

<400> 94
ccgggctcgg gcaagcactt taacctttta agcccaacca gatgagttgc ctgcagtttt 60
ggaggccttc agagcatttc actagacctc tgtctgtgtc ggtccagtgt ctttagccaa 120
gctttgatta aagatgactt ccttgtttgc tcaagaaatt cgcctttcta aaagacatga 180
agaaatagta tcacaaagat taatgttact tcaacaaatg gagaataaat tgggtgatca 240
acacacagaa aaggcatctc aactccaaac tgttgagact gcttttaaaa ggaaccttag 300
tcttttaaag gatatagaag cagcagaaaa gtcactacag accaggattc acccacttcc 360
acggcctgag gtggtttctc ttgagactcg ttactgggca tcagtagaag aatatattcc 420
caaatgggaa cagtttcttt t 441

<210> 95
<211> 441
<212> DNA
<213> Homo sapien

<400> 95
gcagagcccc cagctcacga cggtcacggc cccaaggcct ggtgccctgc tgacgggcac 60
caccgtggcc aacctccagg tggcccggct cacccgggtt cccacttctc agctgcaggc 120
gcaagggcag atgcagaccc aggcacccca gccagcccag gtggccttgg cgaagcctcc 180
ggtggtgtcc gtcccggcag ctgtggtctc ctcaccggga gtcaccaccc tgcccatgaa 240
cgtcgcgggg atcagcgtgg cgatcggtca gccacagaag gcagcaggac agaccgtggt 300
ggcccagccc gtgcacatgc agcagctgct gaagctgaag cagcaggccg tccagcagca 360
gaaggccatc cagccccagg ctgcacaggg cccggcagcc gtccagcaga agatcaccgc 420
acagcagatc accacccctg g 441

<210> 96
<211> 441
<212> DNA
<213> Homo sapien

<400> 96
gcctcagtcc tgtattaggg tacagtggtg gccccagaaa ggaggtaatc aagtctgctt 60
aggggtggta ggattaggaa aggcttcatg tttagggagg aactagcact tgaacttggg 120
tctttaaaca gtcattcctc taaagcaact cacatagaag caactggagc agctagatag 180
tctgatcctt tggatcaggt gcccagaaac cactgatggg cagcagggaa gaggagatgc 240
agtgagcttg gctgttggga agaaattcag cctgcttaag ttccaactga atgaaagatg 300
ctcagggcaa tgctgtgagg aagaatgtgc tttaaacaat gacagcctat actgtctgca 360
gggcgtgcgt gctgcacact gtatgctggt atttggaagt catgctgcca ttctgtactg 420
tctctcctgt gtttgagagc t 441

<210> 97
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 97
cgccctcgcc cccgcccccc tttgngccgg caaaactttg gcactggggc acttggcagc 60
gcggggagcc cgtcggtaat tttaatattt tattatatat atatatctat atttttgtcc 120
aaaccaaccg cacatgcaga tggggctccc gcccgtggtg ttatttaaag aagaaacgtc 180
tatgtgtaca gatgaatgat aaactctctg cttctccctc tgcccctctc caggcgccgg 240
cgggcgggcc ggtttcgaag ttgatgcaat cggtttaaac atggctgaac gcgtgtgtac 300
acgggactga cgcaacccac gtgtaactgt cagccgggcc ctgagtaatc gcttaaagat 360
gttcctacgg gcttgnngct gcngatggtt tgtnttgntc tgtttcttgg tctttttttng 420
nattataaaa aataaccnat g 441

<210> 98
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 98
agccgaggcc acagtttgct attttnttat gaaaggagga tctgtttggg aaacatagat 60
tgtcttcccc tcaaatgagg ggaaaaaaaa agacccttg ttcaaatgga ttctgttgta 120
aaaaattatt tttaaaggaa atcacaaatt gtatgtcatt cttaatgcta gtcttataga 180
ataaatccat aaaattgttt ttatgttcag tatgtttatg tcattctaaa tgcagcaaat 240
tcaatgatag cagttcaatt gactcatagc agtgttttgt atttttcta attctttagc 300
tttcaatatt ggattaaagt cttgtttgtg aatatagttt ccgtatggca aatgatttct 360
tgcttattag cttttgttaa agaatgctta gtaagagcta agcttttaaa agtaatgcaa 420
acatttatcg ttaataaaac c 441

<210> 99
<211> 441
<212> DNA
<213> Homo sapien

<400> 99
agccgggagc agcagaggtc tagcagccgg gcgccgcggg ccgggggcct gaggaggcca 60
caggacgggc gtcttcccgg ctagtggagc ccggcgcggg gcccgctgcg gccgcaccgt 120
gaggggagga ggccgaggag gacgcagcgc cggctgccgg cgggaggaag cgctccacca 180
gggcccccga cggcactcgt ttaaccacat ccgcgcctct gctggaaacg cttgctggcg 240
cctgtcaccg gttccctcca ttttgaaagg gaaaaaggct ctccccaccc attcccctgc 300
ccctaggagc tggagccgga ggagccgcgc tcatggcgtt cagcccgtgg cagatcctgt 360
cccccgtgca gtgggcgaaa tggacgtggt ctgcggtacg cggcggggcc gccggcgagg 420
acgaggctgg cgggcccgag g 441

<210> 100
<211> 431
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(431)
<223> n = A,T,C or G

<400> 100
cgaattcaag ccagcctggg gctgaggctg aggcactggc gaggagaggg cgctcctctc 60
tgcacaccta ctagtcacca gagactttag ggggtgggat tccactcgtg tgtttctatt 120
ttttgaaaag cagacatttt aaaaaatggt cacgtttggt gcttctcaga tttctgagga 180
aattgctttg tattgtatat tacaatgatc accgactgaa aatattgttt tacaatagtt 240
ctgtggggct gttttttttgt tattaaacaa ataatttaga tggtgaaaaa aaagcttgcg 300
gccgcactcg agcccgggtg aatgattgag tttaaaccgc tgagcaataa ctagcataac 360
cccttggggc ctntaaacgg gtnttgaggg gttttttgnt gaaaggagga actatatccg 420
gataacctgg c 431

<210> 101
<211> 441
<212> DNA
<213> Homo sapien

<400> 101
cctgccacaa gctcatcatc cgacacttag cttctgatgc catcattaat gagaactatg 60
actacctgaa ggggttcttg gaagacctgg cacctccaga gcgcagcagc ctaattcagg 120
attgggaaac atctgggctt gtttacctgg actatattag agtcattgaa atgctccgcc 180
atatacagca ggtggattgc tcaggtaatg acctggagca gttacacatc aaagtgactt 240
cactgtgcag tcggatagag cagattcagt gttacagtgc taaagatcgc ctggctcagt 300
cagacatggc caaacgtgta gccaacctgc tgcgcgtggt gctgagtctg catcatcctc 360
ctgatagaac ctccgactca acaccagacc ctcagcgagt ccctttgcgc ctcttggctc 420
cccacattgg ccggcttccc a 441

<210> 102
<211> 431
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(431)
<223> n = A,T,C or G

<400> 102
ccgaattcaa gcccgggatc tccttgcgct cggtcctcta cgtggagtca cctatgcaga 60
ggaattccac ggggcggggg cgaggacagg gtgcgggggt ctttatggca gacaatcccc 120
ggctgagcgc ttggccagag tttctgtgat gctagaatct ggactgcctg cgacctctcc 180
gggactcgga caccagccct cgcctcctgg tgatctttta ggtcctgcag agaagtgaag 240
aggtattgga cgtggccagg gtcaatagtg tgaagccaga attagaattg aattgaaatg 300
ccttcggttt tgatatctct gctgtttgtc ttgaggcaac tgcttttcct ccttnggncc 360
tctcagtttn cnatataang aggcccaaag gcgaanccac ttgagctcgt gagngacagn 420
cagatgtcaa g 431

<210> 103
<211> 448
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(448)
<223> n = A,T,C or G

<400> 103
gaattcaagc cacatgccta tcatatagta aaacccagnc catgacccct aacaggggcc 60
ctctcagccc tcctaatgac ctccggccta gccatgtgat ttcacttcca ctccataacg 120
ctcctcatac taggcctact aaccaacaca ctaaccatgt accaatgatg gcgcgatgta 180
acacgagaaa gcacatacca aggccaccac acaccacctg tccaaaaagg ccttcgatac 240
gggataatcc tatttattac ctcagaagtt tttttcttcg caggattttt ctgagccttt 300
taccactcca gcctagcccc taccccccaa ttaggagggc actggccccc aacaggcatc 360
accccgctaa atcccctaga agtcccactc ctaaacacat ccgtattact cgcatcagga 420
gtatcaatca cctgagctca ccatagtc 448

<210> 104
<211> 447
<212> DNA
<213> Homo sapien

<400> 104
gaattcttat gtctatagac ttccaatcag aagtctcact ggtggggctg ggggtggggg 60
caggcaggag gcatggatgg gaacctgagt aggtagtgtg gccaagagat cagcacaacc 120
tttgcaggct gacttgctaa gtctgacagt gacaaacttg tgagcttact gcagtcagtc 180
acagaggctg ttctttttca cacacccctt catgcctggc tttccccata tccacatgca 240
gagggcgagc tcataaaact acagggaagc gtgaaatgat ggctttggta gctgtttact 300
gggtaacccc actgtgacac tgtccttttc atgtgatgtg gaaacctact tctgtcctcc 360
aaaccatgaa atgtgtcatc tagactgcag agtacttgag tgctttgcct cccgatatgc 420
cagagcttgt ggtccaaagc ccattcc 447

<210> 105
<211> 447
<212> DNA
<213> Homo sapien

<400> 105
gaattcaagc gaggagcgct cgttctagtt cgtccaccat ggcgtccgtg gggaccctcg 60
ccttcgatga gtatgggcgc ccctttctca ttatcaagga ccaagatcgc aagtcccgtc 120
tcatggggct tgaggccctc aagtctcaca tcatggctgc caaagctgta gcaaacacaa 180

tgcggacgtc actgggacca aacgggctgg acaagatgat ggttgataag gatggcgatg 240
tgactataac aaacgatggt gccaccattc taagcatgat ggatgtcgat catcagattg 300
ccaagctgat ggttgaactg tccaaatccc aggatgatga aattggagat gggaccacag 360
gagtggttgt cttggctggg gccttgttgg aagaagctga acagctgctg gaccgaggca 420
tccacccaat cagaattgct gatggct 447

<210> 106
<211> 451
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(451)
<223> n = A,T,C or G

<400> 106
cgaattcnag cagcggtccc acaaagcact ttcttaaacc ttgagaatct ccaagagaaa 60
aatatttggg gaaggaggga ggaaatatgt cccttgcaca ccacccctga agcacatggc 120
agtaggaaac agcataggat tgtatgtggg aggtggatag gtcggtgatg tgtggagcgg 180
aaaagcaggt tggtaaagtt cccttcttgg gacttattcc tggagtcagt ggatacaagt 240
agtgcagagg gttcacactg caaatagtgt tctcatctca aagcaaacta tcattccaga 300
aggaaaagtg tgtcagggca agcagacaac acaatttcct atcagaatat gtcccctcaac 360
ccccgaaaca aggcttctct cagcctcccc accagtgatg gataacagct cctattctca 420
gctgacctga ctgagccaac ccatgaactc t 451

<210> 107
<211> 451
<212> DNA
<213> Homo sapien

<400> 107
ccgaattcaa gcagaaaact gcctttattc tattagtagt tggaaaaatt aactggtaca 60
gaaaaaaagt ttagtcagct ggagagaaga gagactgagt gccacccatg agaactggtg 120
gctcctctgg gagggaacct ggatacagtg aggagaaaag agcactgtga attagagcca 180
gacgcttaag tccaggtgag acaggttatg ccatcttcca aagtgtctaa ttgcctcagg 240
cgtgaaacca attcctattt acttagccca gctccatggg gtactgagat acatggggcc 300
gaaaaggggt aatatggcca tcttttatca gaaaaagtga caaaacggga atttaaaaaa 360
tgaattttcc atctgacttt atttccaaat acactttctt ttttaaaaaa ccaatacact 420
ttctttgagg atgacagtat taggaaatcc a 451

<210> 108
<211> 441
<212> DNA
<213> Homo sapien

<400> 108
cgaattcaag cgtaattttt aactttaaaa aaacaaaaac atgaaatccc tcttaacatg 60
ctactgtatg ttccattcca acaggtcgag gagagcttaa acaccttctt cctctgcctt 120
gtttctcttt tatttttttat ttttttcgcat cagtattaat gtttttgcat actttgcatc 180
tttattcaaa agtgtaaact ttctttgtca atctatggac atgcccatat atgaaggaga 240
tgggtgggtc aaaaagggat atcaaatgaa gtgatagggg tcacaatggg gaaattgaag 300
tggtgcataa cattgccaaa atagtgtgcc actagaaatg gtgtaaaggc tgtcttttttt 360
tttttttttta aagaaaagtt attaccatgt attttgtgag gcaggtttac aacactacaa 420
gtcttgagtt aagaaggaaa g 441

<210> 109
<211> 441
<212> DNA
<213> Homo sapien

<400> 109
ttttggttgg ttggttttat tggcaagcat caatgccgca ttacccgtcg ctttccagtc 60
ttcatcgtgc tgccgtgaga ccactggaaa gcacagcatt gtctgcacaa ggctgggact 120
gctccttcac cacagggtag aagtatgggt gctccatggc ctctttggca gtcagtctct 180
gttgatggtc gtatcgcaga agtttgtcca gaagatctag ggcctcaggg ctgacaaggt 240
gtctgttctc actatggata aagttttccc agcgtttccg tgaatgttgt cccaggatat 300
cgttgaagtg tggatctagg tctatgtgat acttcttcag atacccatac agttcttctg 360
tacccagaac cttggcaatg cgaacaagct ggtcatagtt gtcctgtcca tggaagaatg 420
gttcccttcg aaagatcatg c 441

<210> 110
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 110
aattcaagcc anagcccggg cgcaggcggc ggatggagcg gaacggctag gggtcttgag 60
aagcaatggc cacagaagcc cctgtgaata tagcaccacc tgagtgtagc actgttgtca 120
gcacagcagt tgacagcctc atttggcagc caaactcact aaatatgcac atgataaggc 180
ccaagtccgc caagggacgg acaagaccga gtctgcagaa atcccagggc gtggaggtgt 240
gcgctcatca tataccatct ccgcctccag ccattcccta tgagttgcca agcagccaaa 300
aaccaggagc ctgtgcaccc aaatctccaa accagggagc ttctgatgag atccctgagc 360
tgcagcagca agtacccact ggggcttcct cttctctcaa taagtatcca gtccttcctt 420
ccatcaacag aaagaacctg g 441

<210> 111
<211> 441
<212> DNA
<213> Homo sapien

<400> 111
tccgaattca agcagacacc actgcatgtg gctgctgcca accgggccac caagtgtgct 60
gaggctctgg caccctgtt gagcagcctc aacgtggctg acaggagcgg gcgcagtgcc 120
ctgcaccatg cagtgcatag tgggcatctt gagacggtga acctgctcct caacaaggga 180
gccagcctga atgtctgtga caaaaaggag cggcagcctc tgcattgggc agcttttcta 240
gggcacttgg aggtcctaaa actgctggtg gcacggggag cagacctcgg ctgcaaggac 300
cgcaagggct atgggctgct ccatacagct gctgccagtg gccagattga agtggtgaag 360
tacctgcttc ggatgggagc ggagatcgat gaacccaatg cttttggaaa cacagctttg 420
cacatcgcct gctacctggg c 441

<210> 112
<211> 441
<212> DNA
<213> Homo sapien

<400> 112
ccgaattcaa gcaaagcagc caggaaggac aggctttccc ctgtatatca taggaaactc 60
agggacattt caagttgctg agagttttgt tatagttgtt ttctaaccca gccctccact 120
gccaaaggcc aaaagctcag acagttggca gacgtccagt tagctcatct cactcactct 180
gattctcctg tgccacagga aaagagggcc tggaaagcgc agtgcatgct gggtgcatga 240
agggcagcct gggggacaga ctgttgtggg aacgtcccac tgtcctggcc tggagctagg 300
ccttgctgtt cctcttctct gtgagcctag tggggctgct gcggttctct tgcagtttct 360
ggtggcatct caggggaaca caaagctatg tctattcccc aatataggac ttttatgggc 420
tcggcagtta gctgccatgt a 441

<210> 113
<211> 441
<212> DNA
<213> Homo sapien

<400> 113
aattcaagcc gccgaagaag catcgttaaa gtctctcttc accctgccgt catgtctaag 60
tcagagtctc ctaaagagcc cgaacagctg aggaagctct tcattggagg gttgagcttt 120
gaaacaactg atgagagcct gaggagccat tttgagcaat ggggaacgct cacggactgt 180
gtggtaatga gagatccaaa caccaagcgc tccaggggct ttgggtttgt cacatatgcc 240
actgtggagg aggtggatgc agctatgaat gcaaggccac acaaggtgga tggaagagtt 300
gtggaaccaa agagagctgt ctccagagaa gattctcaaa gaccaggtgc ccacttaact 360
gtgaaaaaga tatttgttgg tggcattaaa gaagacactg aagaacatca cctaagagat 420
tattttgaac agtatggaaa a 441

<210> 114
<211> 441
<212> DNA
<213> Homo sapien

<400> 114
ttatgtgttg tcgtgcaggt agaggcttac tagaagtgtg aaaacgtagg cttggattaa 60
ggcgacagcg atttctagga tagtcagtag aattagaatt gtgaagatga taagtgtaga 120
gggaaggtta atggttgata ttgctagggt ggcgcttcca attaggtgca tgagtaggtg 180
gcctgcagta atgttagcgg ttaggcgtac ggccagggct attggttgaa tgagtaggct 240
gatggtttcg ataataacta gtatggggat aaggggtgta ggtgtgcctt gtggtaagaa 300
gtgggctagg gcatttttaa tcttagagcg aaagcctata atcactgcgc ccgctcataa 360
ggggatggcc atggctaggt ttatagatag ttgggtggtt ggtgtaaatg agtgaggcag 420
gagtccgagg aggttagttg t 441

<210> 115
<211> 441
<212> DNA
<213> Homo sapien

<400> 115
tgatcttagc ccagaagttg caacacatca acccattatt gcctgcctgc cttaacaaag 60
aggagagcaa aacctttgtt tcaagtttca tgtccgaatt gtctccagtc agagcagaac 120
ttcttgggtt ccttactcat gcccttctgg gggatagttt ggctgctgaa taccttatat 180
tacatctcat ctccacagta tatacaagaa gagatgtcct tccactagga aaatttacag 240
ttaacttgag tggttgccca cggaatagta ccttcacaga acacttgtat cgaattattc 300
aacatcttgt tccagcagta agatgaatat aaatatgtat tataaaccta agcacactta 360
aaaaatccat ctaactgtct ttcattcact tcaaatattt aacttagaga ttttaataga 420
tcttatttaa ggggaaacaa a 441

<210> 116
<211> 266
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(266)
<223> n = A,T,C or G

<400> 116
caagcatgct ctccccgcag ggctcnctca tgggcccccc gccccagcag aacctcatgg 60
tgtcccaccc ccttcggcag cgcagtgtgt ccctggacag ccagatgggc tacctcccgg 120
caccaggcgg catggccaac ctgcccttct agaagtcgct gccagggctg gagccggggc 180
aatgttgcaa atacgataac cttaacaaag ttcttccccc caatgttggg atggcctggg 240
tcgtggggtg gggtggaggg ggtggg 266

<210> 117
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 117
caagcggata cccatggctc cgatgnagat ggggatgggg agattgtgga cgaggatgca 60
gcggtggcgg aggcccttgc agctttagaa gctgctactg caggagaaga tttggatgag 120
actgattagg ggaggggatt tgcacaggga ggtaagctgg tgtcatgctg agcatgcaga 180
tgcatttgct ccctggatgc atagcaggtg attctgccag catgcaccag tgcagcctta 240
ccagttgttt acatccagca tctgttctga ttgtcagcat ctgtcccatg ctgcttgtca 300
catatctgga gtttcactct gtgtagatga gctgtcattc aggacactag gagaaaaatc 360
tgagtgggtc attgtgccca tatccacaga aaatgcagaa gttgaacagc ttgcttgaca 420
accctcaaac atctttgagc a 441

<210> 118
<211> 441
<212> DNA
<213> Homo sapien

<400> 118
tgcaaacaga actagtaaca aagggccgag tgccatatcc aggtatggtg aaccgtgaag 60
tactagaaca agtggagcga ggatacagga tgccgtgccc tcagggctgt ccagaatccc 120
tccatgaatt gatgaatctg tgttggaaga aggaccctga tgaaagacca acatttgaat 180
atattcagtc cttcttggaa gactacttca ctgctacaga gccacagtac cagccaggag 240
aaaatttata attcaagtag cctatttttat atgcacaaat ctgccaaaat ataaagaact 300
tgtgtagatt ttctacagga atcaaaaaaa aaaaaaaagc ttgcggccgc actcgagccc 360
gggtgaatga ttgagtttaa accgctgagc aataactagc ataacccctt ggggcctcta 420
aacgggtctt gaggggtttt t 441

<210> 119
<211> 441

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 119
agcgtggaca gggtcctaga ggantggggc ctggaactcc agcaggatat ggtagaggga 60
gagaagagta cgaaggccca aacaaaaaac cccgatttta gatgtgatat ttaggctttc 120
attccagttt gttttgtttt tttgtttaga taccaatctt ttaaattctt gcattttagt 180
aagaaagcta tctttttatg gatgttagca gtttattgac ctaatatttg taaatggtct 240
gtttgggcag gtaaaattat gtaatgcagt gtttggaaca ggagaatttt tttttccttt 300
ttatttcttt attttttctt ttttactgta taatgtccct caagtttatg gcagtgtacc 360
ttgtgccact gaatttccaa agtgtaccaa ttttttttt tttactgtgc ttcaaataaa 420
tagaaaaata gttataatat t 441

<210> 120
<211> 441
<212> DNA
<213> Homo sapien

<400> 120
aagcgcggcg cggaggccgc ggctggggtt gagccgccgg aggccgcccc cggcgggcgg 60
gccttctggc ggccgccccc ggtttctccc ggggggtgat gagcgggagc ggctctgggc 120
cgagctactg cgcacggtga gcccggagct gatcctggat cacgaggtgc cttcactgcc 180
cgccttccca ggacaggagc ccaggtgcgg cccggagccc actgaagtct tcactgtcgg 240
acccaagacc ttttcctgga cacccttcc gccggacctg tggggcccgg gccgttccta 300
ccggctgctt cacggggcag gagggcacct ggaatccccc gccaggtccc tgccccagcg 360
cccggcacct gatccctgca gggcccccag ggtggagcag cagccgtctg tggagggtgc 420
ccgcggccct gcgcagctgc c 441

<210> 121
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 121
acagcccccc atcgacttnc gcgacgtgga catcggcgag ctgagcagng acgtcatctc 60
caacatcgag accttcgatg tcaacgagtt tgaccagtac ctgccgccca acggccaccc 120
gggggtgccg gccacgcacg gccaggtcac ctacacgggc agctacggca tcagcagcac 180
cgcggccacc ccggcgagcg cgggccacgt gtggatgtcc aagcagcagg cgccgccgcc 240
acccccgcag cagcccccac aggccccgcc ggccccgcag gcgcccccgc agccgcaggc 300
ggcgccccca cagcagccgg cggcaccccc gcagcagcca caggcgcaca cgctgaccac 360
gctgagcagc gagccgggcc agtcccagcg aacgcacatc aagacggagc agntgagccc 420
cagccactac agcgagcagc a 441

<210> 122
<211> 441

<212> DNA
<213> Homo sapien

<400> 122
gaaatagtac cttcaatact taaaaatagt cttccacaaa aaatacttta tttctgatct 60
atacaaattt tcagaaggtt attttcttta tcattgctaa actgatgact taccatggga 120
tggggtccag tcccatgacc ttggggtaca attgtaaacc tagagtttta tcaactttgg 180
tgaacagttt tggcataata gtcaatttct acttctggaa gtcatctcat tccactgttg 240
gtattatata attcaaggag aatatgataa aacactgccc tcttgtggtg cattgaaaga 300
agagatgaga aatgatgaaa aggttgcctg aaaaatggga gacagcctct tacttgccaa 360
gaaaatgaag ggattggacc gagctggaaa acctccttta ccagatgctg actggcactg 420
gtggtttttg ctctcgacag t 441

<210> 123
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 123
atgccaaggc tctgtgggag gatgaaggag tgcgtgcctg ctacgaacgc tccaacgagt 60
accagctgat tgactgtgcc cagtacttcc tggacaagat cgacgtgatc aagcaggctg 120
actatgtgcc gagcgatcag gacctgcttc gctgccgtgt cctgacttct ggaatctttg 180
agaccaagtt ccaggtggac aaagtcaact tccacatgtt tgacgtgggt ggccagcgcg 240
atgaacgccg caagtggatc cagtgcttca acgatgtgac tgccatcatc ttcgtggngg 300
ccagcagcag ctacaacatg gtcatccggg aggacaacca gaccaaccgc ctgcaggagg 360
ctctgaacct cttcaagagc atctggaaca acagatggct gcgcaccatc tctgtgatcc 420
tgttcctcaa caagcaagat c 441

<210> 124
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 124
aaattaaaaa agaggccact gctatttgaa agagttgctc agaaaaatgc aagaatggca 60
gcagaaaagc attattctaa taccctaaaa gcactaggaa tatctgatga gtttgtttca 120
aanaaaggcc aaagnggaaa agtccttgag tncttcanca atcaananac gaaaagngtc 180
actgaanaca aanaaagctt taatgaanaa naaaaaatag aanaaagaga gaatggggaa 240
naaaattatt ttattgatnc cancagccag gattcttnca aggaaaaana tgaagccaat 300
gaggaaagtg aanaananaa atctgttgaa naatcncact gaatcatcaa ggtctcctct 360
ctatgccctt gctgttgttt gcagcgtcag ggngtcagca gccgcatttg ngtttanaac 420
atctgtgggg acgcttntga t 441

<210> 125
<211> 426

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(426)
<223> n = A,T,C or G

<400> 125
gaattntcat tgagaactaa cacttcttag atagngcgtt ttcctctacc agacactgtt 60
gcatgccttg tagggtttgc tcttcctggc agcgttgtga agtagatgga actgtgctta 120
tttgaaggct gggacagaag agattccagg atgttaagtg aattgcctgg ggatggtgct 180
gggacagaca ggtcacatgt ccacatctag cagctgtggg tctctttgtg gtcaccctcc 240
cagtttggcc ctgtgaatgg tcagatctgt aaaggcggat ttctgggaca ttgctcagct 300
gaaacctcct tccttcccca ggctctacgg tttctccana anaattcctc taagtttcat 360
ttcanacnca ccaggatgtt gccggttagt ggcgcattcc acacccgcct catggagcca 420
gccgtg 426

<210> 126
<211> 441
<212> DNA
<213> Homo sapien

<400> 126
catttataag atccttctcc catgagaatg ctgtgttggt taacaaggac tgatcttaaa 60
cctaagcttc tcccaaatcc attccattta taatctctgc taagggtggt ggcagatttg 120
tgaatggttg acatttgttt caggtgtctc tcctgtttta tctgctgcat ctctaactgg 180
ccatcacaac cgcaggctgc tttcaacttc gagaactgca gcacatcttg aatgtgtttt 240
tccactgata ctacctggaa taattcttct ttggaaattc cccaacttgg cattgattcc 300
ttggatttag accttctctt ccagtctgag tgggctcctg ttgagatttc tctctccacc 360
atgaagggct cacccccttc ctccaactgg cagatcacat atggtttgga tactagaagg 420
cccagaatgg ccaagttcct g 441

<210> 127
<211> 441
<212> DNA
<213> Homo sapien

<400> 127
gaccgggggc ggggctccga ggcccgggcg caaccacggg ctcccaggca gcctccgcca 60
gccggacccc gtcgccctcc tgatgctgct cgtggacgct gatcagccgg agcccatgcg 120
cagcggggcg cgcgagctcg cgctcttcct gacccccgag cctggggccg aggcgaagga 180
ggtggaggag accatcgagg gcatgctcct caggctggaa gagttttgca gcctggctga 240
cctggtgagt ggctgcctgg aaggcgtggg tttaggccca ggccagactt caggcccttg 300
ctggagtttt ctgcggaagc cattttggtg atcaggagtg atacttcaca gatcctggag 360
gaaaacatcc cagtccttaa ggccaaactg acagaaatgc gtggcatcta tgccaaagtg 420
gaccggctag aggccttcgt c 441

<210> 128
<211> 422
<212> DNA
<213> Homo sapien

<400> 128
ttttcgagga cttctgcggg agttgcgcta cctgagcgcg gccaccggcc gaccctatcg 60

cgacaccgcg gcctatcggt accttgtgaa ggctttccgt gcacatcggg tcaccagtga 120
aaagttgtgc agagcccaac atgagcttca tttccaagct gccacctatc tctgcctcct 180
gcgtagcatc cggaaacatg tggccctaca tcaggaattt catggcaagg gtgagcgctc 240
ggtggaggag tctgctggct tggtgggtct caagttgccc catcagcctg gagggaaggg 300
ctgggagcca tgaacatgga gaatatcctt ggatgctgca ttcataggag aattgaataa 360
tttctatcaa tatgtattta tcattaaatt ttttttaagt ttagcttgcg gccgcactcg 420
ag 422

<210> 129
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 129
cagaggcagc tccggcggcc gagaggaggg agcgcggcgc agagaggagg ggcttgcgcc 60
ccgtagaaat gtcaatcaga gcccggagcc ccgggaatct ccgccaatct gttcggacct 120
gacctggctc ctcgccgccg ccctcgccgc cgccgtcgcc gccgcggagc agatcaatag 180
gcgaacgcgg agcacagcgc agcgcgggcg gcagcgcggc cccaagcccg gcccagcccc 240
cgatgcgcgc cggagcccgc gggcggcgct gagctgggcg gcccgggggt cgggcccccc 300
ctccgtccgt gccccgcggg ccaccatgtc cttccccccag ctgggatacc aatacatccg 360
cccgctttac ccgtccgagc gcccgggggc cgctggcggc agcggcggca gcgcgggggc 420
ccggggcggn ctgggtgccg g 441

<210> 130
<211> 441
<212> DNA
<213> Homo sapien

<400> 130
acgtaaaacc tggccttttg tagacgtctg acgattagtt tttgaaatac tttccgtttt 60
ctgtaaaatc aaagaaggaa aaagtttcag gtactcttga cactcctgag aagactgtgg 120
atagccaggg ccccacacca gtttgtacac caacattttt ggagaggcga aaatctcaag 180
tggctgaact gaatgatgat gataaagatg atgaaatagt tttcaaacag cccatatcct 240
gtgtaaaaga agaaatacaa gagactcaaa cacctacaca ttcacggaaa aaaagacgaa 300
gaagcagcaa tcagtgattt tcaatgtatt atatttcttt tgaaaaatat aatatttttа 360
tgagagtgga ctttgtattt cactaggtac aatggaatac aacctttgac aagattttca 420
gaggaaaaat acactgtttg g 441

<210> 131
<211> 441
<212> DNA
<213> Homo sapien

<400> 131
ctgcaacagc cgctgcactg ccactcagtt ttctaaggaa ctcctcctac taccatcttg 60
gctcagtctc cctcacttaa gccctgggtt tgaaaaatta attgcaactt cccaggaaac 120
attgttcagt ttgcagatta agcctggcac tcacctatca gaaaccagag ctccgcctgc 180
ttagttgttt caaagttttc tgaaagaaaa ctaggggagc acttgtgaac acaggagcag 240
ctggtgatct gctttcttac cctaactctt gacaaatgag tcgtctacta ttttaaagag 300
tctggaggtc tctgactctg ccataacaat aacctgctgt taatttataa cacagatttt 360

tgtttggaag agccttattt gaaatacact ttgatttatt ttcttaaata tttatattct 420
tttcttgctt acttcagggt t 441

<210> 132
<211> 441
<212> DNA
<213> Homo sapien

<400> 132
gtggtaatgc cagccacact cctcagagcc gtggccagat ctcatcatat attatcaaaa 60
gcacatcagt gccgaagaat cggtcatcta atgttaaaac cacttaagga atttgaaaat 120
acaacatgca tcacactgac aatacgtcaa agcttggatt tgttccttcc tgataaaaca 180
gctagtggtt tgaataagtc tcagatcctg gaaatgaacc aaaaaaagtc agataccagc 240
atgctgtctc cattaaatgc tgctcgttgc caagatgaaa aggcacacct tccaaccatg 300
aaatcctttg gtactcacag gagagtgacc cacaaaccaa atctgttggg ttctaaatgg 360
tttataaaaa tattaaagag gcatttctca tctgtatcaa cggaaacatt tgttccaaaa 420
caagacttcc cacaggtgaa g 441

<210> 133
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 133
catgacacag caagacgana agaccctatg gagctttaat ttattaatgc aaacagtacc 60
taacaaaccc acaggtccta aactaccaaa cctgcattaa aaatttcggt tggggcgacc 120
tcggagcaga acccaacctc cgagcagtac atgctaagac ttcaccagtc aaagcgaact 180
actatactca attgatccaa taacttgacc aacggaacaa gttaccctag ggataacagc 240
gcaatcctat tctagagtcc atatcaacaa tagggtttac gacctcgatg ttggatcagg 300
acatcccgat ggtgcagccg ctattaaagg ttcgtttgtt caacgattaa agtcctacgt 360
gatctgagtt cagaccggag taatccaggt nggtttctat ctacttcaaa ttcctccctg 420
tacgaaagga caagagaaat a 441

<210> 134
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 134
ctttattgga attattttga taaaagcaaa tacttgtata taagacaaat ataggaaata 60
gaagctatct acttgaagtg ccccctaatt ttcaggcatt tttcccccaa taaggngaca 120
gctgctcaca caggataata caggaaagtt gtaagttctt aggatatgcc cagcaccagg 180
cggaanatat tgtgtcagca aattccatag caaggaanga gaacaggcat ggctctccac 240
ctctcccagc ctgaaggnca agttgcagga atcactgtct aagcatgaca gccatacaaa 300
gcttctctcc ttgaggggga gattccaaac cttagacctt cactcagatt agtgcanaat 360

ctctgtgaag gtgcattgnc cctggctggg cttccagagt tnattnttta ctcnnggata 420
acacggcgan naacntctna c 441

<210> 135
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 135
agacaacctt agccaaacca tttncccaaa taaagtatag gcgatagaaa ttgaatcctg 60
gcgcaataga tatagtaccg caagggaaag atgaaaaatt ataaccaagc ataatatagc 120
aaggactaac ccctatacct tctgcataat gaattaacta gaaataactt tgcaaggaga 180
gccaaagcta agacccccga aaccagacga gctacctaag aacagctaaa agagcacacc 240
cgtctatgta gcaaaatagt gggaagattt ataggtagag gcgacaaacc taccgagcct 300
ggtgatagct ggttgtccaa gatagaatct tagttcaact ttaaatttgc ccacagaacc 360
ctctaaatcc ccttgtaaat ttaactgtta gtccaaagag gaacagctct ttggacacta 420
ggaaaaaacc ttgtagagag a 441

<210> 136
<211> 425
<212> DNA
<213> Homo sapien

<400> 136
tgcttcacgc ggaaatacac gctgcccccc ggtgtggacc ccacccaagt ttcctcctcc 60
ctgtcccctg agggcacact gaccgtggag gcccccatgc ccaagctagc cacgcagtcc 120
aacgagatca ccatcccagt caccttcgag tcgcgggccc agcttggggg cccagaagct 180
gcaaaatccg atgagactgc cgccaagtaa agccttagcc cggatgccca cccctgctgc 240
cgccactggc tgtgcctccc ccgccacctg tgtgttcttt tgatacattt atcttctgtt 300
tttctcaaat aaagttcaaa gcaaccacct gcttgcggcc gcactcgagc ccgggtgaat 360
gattgagttt aaaccgctga gcaataacta gcataacccc ttggggcctc taaacgggtc 420
ttgag 425

<210> 137
<211> 441
<212> DNA
<213> Homo sapien

<400> 137
agagcttatt ggcacttagc cattcattgg tcctgatgga gttaagtgag acagcttacc 60
tcatctatca agtgacactc atttccccac tcctaggata ccctttctga ggggctacat 120
ccttccaagt gtttacaatc tagtctcaaa actttagtgt tctctgtgag tgccaggttc 180
attttagggt gagatatcat agactatgtt atttagctac cataccgaaa taggtatgta 240
acatattttg gtgattttcc aaatagcata caaatgtaac attttggtgg ttttccaaat 300
agcagttttc aaaaatattt gctttagtgg ttaatatatg attctcttgt gtctctgtta 360
tcaataatgg gcatgataaa aaatccagaa tatgagagat attggcactc tgaggatcat 420
cttctgaatt tgaaaaggat t 441

<210> 138
<211> 441

<212> DNA
<213> Homo sapien

<400> 138
tctgcaacaa tgccacatgg gcaattggag aaatctccat tcaaatgggt atagagatgc 60
agccttatat tcctatggtg ttgcaccagc ttgtagaaat cattaacaga cccaacacac 120
caaagacgtt gttagagaat acagcaataa caattggtcg tcttggttac gtttgtcctc 180
aagaggtggc ccccatgcta cagcagttta taagaccctg gtgcacctct ctgagaaaca 240
taagagacaa tgaggaaaag gattcagcat tccgtggaat ttgtaccatg atcagtgtga 300
atcccagtgg cgtaatccaa gattttatat ttttttgtga tgccgttgca tcatggatta 360
acccaaaaga tgatctcaga gacatgttct gtaagatcct tcatggattt aaaaatcaag 420
ttggcgatga aaattggagg c 441

<210> 139
<211> 441
<212> DNA
<213> Homo sapien

<400> 139
attttccctg ttctatctca aaaatgtaag aatttgcatt caatttccca actttgtctg 60
ggtttgtttc aatgcagaaa gggttttgtt tactttggtc tggtactttg agtaaaaatg 120
acctagtcag gtgcttcaaa tttatgctgc agctttgata tcacagacac acggtctcct 180
gagggtttag acccatcatt ttggaaggca tggccatagg cagtgtgaga ggcaggggtc 240
actcatggag tattgtagga ggctcagacg gctgcaacgg ctgctcgtgc atagaccaaa 300
tggatggtgt agggagagac cgggagagga gaccatggat caggctctca cagtggttta 360
gggatgaggg gataatggcc tgatttagga tgggggagtg acatgggagc aagagaacga 420
tatgaaagtt ccaacactgg t 441

<210> 140
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 140
tcatttattt acatattata tgtccgatat ataatgtgta gataaatgta tactgcagtg 60
aaagtgacca ctctactgaa ctgtacagca cattatagga caaattattg gaagttttca 120
tctcatactg gtatctttta ataaaaaaaa aaaattaaaa atcaaagaaa acgtctggcc 180
agggccagga taagacaaac aggaacactt cagaatccaa ggcaggggaa ggctgttggc 240
ctctcttcan aggactgcag gggtgggaag agggagggac agatcatgca caaaagtact 300
tacaaattac acacccaacc cccaccctca ataaaaacag aagaaaaagc cccatcactt 360
aacaccaaac agcaacatta tcaatacacc ctttctttgc tgcccaccat gccttatttg 420
aaagaggagg cctgtgagaa g 441

<210> 141
<211> 441
<212> DNA
<213> Homo sapien

<400> 141
agcaattgca gttaagtaag ttacactaca gttctcacaa gagcctgtga ggggatgtca 60

ggtgcatcat tacattgggt gtctcttttc ctagatttat gcttttggga tacagaccta 120
tgtttacaat ataataaata ttattgctat cttttaaaga tataataata ggatgtaaac 180
ttgaccacaa ctactgtttt tttgaaatac atgattcatg gtttacatgt gtcaaggtga 240
aatctgagtt ggcttttaca gatagttgac tttctatctt ttggcattct ttggtgtgta 300
gaattactgt aatacttctg caatcaactg aaaactagag cctttaaatg atttcaattc 360
cacagaaaga aagtgagctt gaacatagga tgagctttag aaagaaaatt gatcaagcag 420
atgtttaatt ggaattgatt a 441

<210> 142
<211> 541
<212> DNA
<213> Homo sapien

<400> 142
agcatcaaga ggaactattt cttagcccac atattcatgt gtcatagttc aggaacacaa 60
gtcagtgaca aacttctagg taattcaacc tgaagaaatt ctttatattc caagattact 120
ttacactctg aaagatacca gccttcctca tctcctcaaa atctttcatg acatcgtagt 180
ttctgtagaa atctgcgtat gccttctttc tttgatcagc cacacgaaac ttatacaaag 240
ctgcaacccc cagggatagc acgaatgcta cagccatatg atttcgcaga cgcctggcca 300
gaaggccacg catccgaggt tttggcaaaa cttcgggagc catggtagtt actgtccttg 360
atacgtatgc taaccccggt ccgtggtccg cccgccgccc tccctcaggc cgcagcggaa 420
gcgtgagaga gtagatgcgg gagaggcctg aagctgcacc acggcggaga cacacagtca 480
cgactaaatc cgaggcagag agaaagcggg agggtgccgt tgtgacaacc cgcagtgtcg 540
g 541

<210> 143
<211> 441
<212> DNA
<213> Homo sapien

<400> 143
ggccccgtcc ccgcgctccc gccgctcccg ccagggcagc ccgggaggcc agacgctggc 60
gctgcaggga gagggcggtg ggcgcatccg ctaggggggcg cggcggggcg gggcgcacct 120
ttcggcgggc ctcgcggatg gcggcgcagg gcgtagggcc tgggccgggg tcggcggcgc 180
ccccggggct ggaggcggcc cggcagaagc tggcgctgcg gcgcaagaag gtgctgagca 240
ccgaggagat ggagctgtac gagctggctc aggcggcggg cggcggtatc gaccccgacg 300
tgttcaagat cctggtggac ctgctgaagc tgaacgtggc ccccctcgcc gtcttccaga 360
tgctcaagtc catgtgtgcc cgggcagagg ctagcgagcg agccccagga ccctgcggcc 420
gtgtctctgc ccacgtcgag c 441

<210> 144
<211> 441
<212> DNA
<213> Homo sapien

<400> 144
ctggcacgtc tgagaatggt ggatgtggtg gagaaagaag atgtgaatga agccatcagg 60
ctaatggaga tgtcaaagga ctctcttcta ggagacaagg ggcagacagc taggactcag 120
agaccagcag atgtgatatt tgccaccgtc cgtgaactgg tctcaggggg ccgaagtgtc 180
cggttctctg aggcagagca gcgctgtgta tctcgtggct tcacacccgc ccagttccag 240
gcggctctgg atgaatatga ggagctcaat gtctggcagg tcaatgcttc ccggacacgg 300
atcacttttg tctgattcca gcctgcttgc aaccctgggg tcctcttgtt ccctgctggc 360
ctgccccttg ggaaggggca gtgatgcctt tgaggggaag gaggagcccc tctttctccc 420
atgctgcact tactcctttt g 441

<210> 145
<211> 426
<212> DNA
<213> Homo sapien

<400> 145
gccaactcga aaaagccgta caaaaaataa gcaaaagcgt cccaggagcc gtactctgac 60
agctgtgcac gatgccatcc ttgaggactt ggtcttccca agcgaaattg tgggcaagag 120
aatccgcgtc aaactagatg gcagccggct cataaaggtt catttggaca aagcacagca 180
gaacaatgtg gaacacaagg ttgaaacttt ttctggtgtc tataagaagc tcacgggcaa 240
ggatgttaat tttgaattcc cagagtttca attgtaaaca aaaatgacta aataaaaagt 300
atatattcac agtgcttgcg gccgcactcg agcccgggtg aatgattgag tttaaaccgc 360
tgagcaataa ctagcataac cccttggggc ctctaaacgg gtcttgaggg gttttttgct 420
gaaagg 426

<210> 146
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 146
ggaacaaaga gaggntgnta ccatttttac tgaaactatt ctaaacaact gaaaaggagg 60
gattcccccc taactcattt tatgaggcca acatcaccct gataccaaaa cctggcagaa 120
atacaataaa aaaagaaaac ttcagaccag tatccctgat gaacatcagt gtgaaaatcc 180
tcagtaacgt actggcaaac caaatccagc agcacatcaa aaagcttatc cagcatgatc 240
aagtaggctt catccctggg atgcaaggct ggttcaacat atacaaatca ataaatgtaa 300
tccatcacgt aaacagaacc aaagacaaaa accacatgat tatctcaata gatgcagaaa 360
aggccttcga taaaattcaa catcccttca tgttaaaaac tctcaataaa ctagatatga 420
atggaacata actcaaaata a 441

<210> 147
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 147
tcccgggacg actgnngctg tgcgcgctcc atgcacgagt tttccgccaa ggacatcgac 60
gggcacatgg ttaacctgga caagtaccgg ggcttcgtgt gcatcgtcac caacgtggcc 120
tcccagtgag gcaagaccga agtaaactac actcagctcg tcgacctgca cgcccgatac 180
gctgagtgtg gtttgcggat cctggccttc ccgtgtaacc agttcgggaa gcaggagcca 240
gggagtaacg aagagatcaa agagttcgcc gcgggctaca acgtcaaatt cgatatgttc 300
agcaagatct gcgtgaacgg ggacgacgcc cacccgctgt ggaagtggat gaagatccaa 360
cccaagggca agggcatcct gggaaatgcc atcaagtgga acttcaccaa gttcctcatc 420
gacaagaacg gctgcgtggt g 441

<210> 148
<211> 541
<212> DNA
<213> Homo sapien

<400> 148
aaccaagctt cagcccactt tcacacccac tgggcaataa actttccatt tccattctcc 60
tagctgggga tggggcatgg tcaaacttag ccatcccctc ctcagcaagg catctaccgg 120
cccctcacag agacagtact ttgaaactca tgttgagatt ttaccctctc ctccaaccat 180
tttgggaaaa ttatggactg ggactcttca gaaattctgt cttttcttct ggaagaaaat 240
gtccctccct taccccccatc cttaactttg tatcctggct tataacaggc catccatttt 300
tgtagcacac ttttcaaaaa caattatata ccctggtccc atctttctag ggcctggatc 360
tgcttataga gcaggaagaa taaagccacc aacttttacc tagcccggct aatcatggaa 420
gtgtgtccag gcttcaagta acttgagttt taatttttt ttttcttggc agagtaattt 480
aaaatttaaa tggggaaaga tatttaatat ttaatactaa gctttaaaaa gaaacctgct 540
a 541

<210> 149
<211> 441
<212> DNA
<213> Homo sapien

<400> 149
aaatggtggg acaataaaat gagttacatt gccacctgag aaacctcaga ggggaggacc 60
cagccttagc ctcccctcctc ccaagtgcaa aatgtgtaaa cagagtaaac ggaacagaaa 120
agtgcagtct aagtggtttt ctctcctgcc cctcccaccg cccctccccc caccccctat 180
tatttgggga taaagaatat aaagacaacc ctggcttttc tattgccttg ttgcttgctg 240
aatataagga atggggtggg gcaggaaggg gcttgccctt agccacagct ctacggctgt 300
gcctcattca tttccacagc tgccagtgtc cctagagttt atcaggtgaa ttggtcaggg 360
gatcagtctc cctcgagcct gacttacggc tgggacagcc ccatctttct gttgattatg 420
tggcgcatat atatatatat a 441

<210> 150
<211> 441
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(441)
<223> n = A,T,C or G

<400> 150
attaggattt atatntgtca ttttaaaaac agtcgattta aaccatgtag aaataagtat 60
gcaaaaagtc tgcaaaacaa aacatacttt aaacatgttt aagtagatag attatctgaa 120
attctggatt tttcagtcac ttcattgtta tgctatggca gccaagtaat ccttaacttc 180
agttggagta agcctcctaa atccagcttc attgcagatt ccaacttcta tgttatcctc 240
tgtcatttgc ccttcaaagc tttcctttag ggttaagatg gctgtatgaa tggcatcttc 300
aagttccaga tcttcattat atcttttctc aaggaaagtc ttcccattca catagttctt 360
tcccattgct gtagctttcc aggcaaagta agctccagat ggatctgact gaaataaata 420
tggtcgtccc tcattccaac c 441

<210> 151
<211> 441
<212> DNA

<213> Homo sapien

<400> 151
aataaatttt tatttgcgtc ttttcaattc tacaaatttt aaaagtcatt atacattttt 60
aaaggttact aaaaatcact gaaatatgtc tccacttaaa atggaatcaa agtaatttat 120
acaattttca cctttaatat tcataattta tttaaaagta gaataaattt taattctaaa 180
atgattatta catcttttta gaactgaaat tacatttcaa tgagtgtagc tccactgtag 240
ttcataaggt ttatacttaa gaaatattta attaacaaat taatttatac caaacaaaaa 300
caaactaaaa aaatgcatac gtataatgag acagtttatg tatggtaatt acagatttga 360
aattaagtct cctttgcaaa gatcagtgcg tctggactgg ttaatccact aaattatcat 420
ctttgaataa gaaaccaaaa a 441

<210> 152
<211> 541
<212> DNA
<213> Homo sapien

<400> 152
tatgttccct aataatacag gtatcacgac ctccctctat taagaactaa tgagttacat 60
aatcctatct acatttccaa attaatcgtc ttatactttt acaataataa tatgccagat 120
tactaataag gaaaacaact gacctctctg actccttcaa gctctccata cctccgtatg 180
tcagaataaa gctattagaa tgcttttctc accaagatga taaagtttca ttaatattct 240
agaaatcatc aaatgggaat ccatgagaat gatgtgtgtc aaggtaggag acataatggg 300
aaaattcagt attatttaag tgtaattttg acccccagga ttccccacca cccctccttg 360
accttaaaaa tcaacaggtc acaagccgcc tgtaatgtgt tgcttttggt gttaattcat 420
cacttttcta ctctgctttt ccacaaaata agtgaatatt catggacaat gtgcttcaaa 480
gatccaagga attcattcca cattggaccc agaggccagg atctcacagt gaaaccacct 540
t 541

<210> 153
<211> 474
<212> DNA
<213> Homo sapien

<400> 153
actcagctat atttagcaac actccatgta gctaatattt tttggtagca tctggtagac 60
cttagaatgt tacatagcca gtaggttctt tattcaaatt ttaagtatct taagaatagt 120
agggcagtaa cagttacttt tgagagtttt ctggtcaagc ttttaccagg cattctctag 180
ccttggtaca aaaaaaaaaa aacctgctgg ttgcgcagat acctaggctt gtccatttta 240
tgcatttcag caaagtcatt ggatactatt gcaacttggg aatactggtc tgcatcaagt 300
ttattcggta gtttgaccgc tagtatgttg gaagttattt ggattgtttt tggaattttg 360
actggctgaa ttatggttgg tataaagtta tgtgtataac tggcaggctt atttatctgt 420
tgcacttggt tagctttaat tgttctgtat tatttaaaga taagtttact caac 474

<210> 154
<211> 530
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(530)
<223> n = A,T,C or G

<400> 154

agcaatcatt tttgtgtgga gcatatttcc ttactcttaa tataggaggg gcagagaagg 60
gagcactggg ggagccagct aatgtccaac agaagcagct ttgacgcagg aaaaggttta 120
ttttgcttta taggctcatt gttattctgt gcactggggc atacatatct tttctaaaaa 180
atcacttaca actgtcagta ttaccagctg gctctcaaag tttaatatca ttgatggtgt 240
gtgaggtaca gttcaggggc actacaaaat ctgtgtacat tgtcttatat acctcctgac 300
cacacctata tctgtagtat ttcttgtgtt agttttgaac atatgtatca aaaaaggaaa 360
gctttatcgt atatgatttt atctagtcct tatttttcca attcacttct gtaagacctt 420
atgtaataaa aacacntgca ttttaaaaag taacagaaat atactaaaaa ttacatgtat 480
aatgaacagt nttaaacatt gagaaaggaa acttaatctc tgtgtccaca 530

<210> 155
<211> 551
<212> DNA
<213> Homo sapien

<400> 155
cattttataa tctttttaat tttaatgcaa gtacactggt gtttatattt gcacagagta 60
ttgatatgtg atgtattaag tcacaaaagt aagctgtgac attgtctata agcatttggc 120
tccacaaatg tatttggatt gttttctatg tgaagcaaac caattataat taaccacatg 180
ttgtagtaac tggtcttttt atatttaagc agaatcctgt aagattgctt gtctttgctt 240
aaaaacaata cctttgaaca tttttgaatc acagaatagc ggtaccatga tagaatactg 300
caattgtggt cagaattaca gtatgcacaa agaattaatt agcattatta aagagtcctc 360
actaaacatt tcatatgatc acactgaaga actgtaacat tccatagagt gaagtggttc 420
aaatttctct tggaattttt acttttgttg gccttatttt atgatccttt tcatatttct 480
tttgacttag agtattaata catggccaaa ataatttagt tactacctca tacaaacaat 540
ataatggtta c 551

<210> 156
<211> 535
<212> DNA
<213> Homo sapien

<400> 156
caaggctcga gatcttcgcg ggaagaagaa ggaggagctg ctgaaacagc tggacgacct 60
gaaggtggag ctgtcccagc tgcgcgtcgc caaagtgaca ggcggtgcgg cctccaagct 120
ctctaagatc cgagtcgtcc ggaaatccat tgcccgtgtt ctcacagtta ttaaccagac 180
tcagaaagaa aacctcagga aattctacaa gggcaagaag tacaagcccc tggacctgcg 240
gcctaagaag acacgtgcca tgcgccgccg gctcaacaag cacgaggaga acctgaagac 300
caagaagcag cagcggaagg agcggctgta cccgctgcgg aagtacgcgg tcaaggcctg 360
aggggcgcat tgtcaataaa gcacagctgg ctgaggcttg cggccgcact cgagcccggg 420
tgaatgattg agtttaaacc gctgagcaat aactagcata acccccttggg gcctctaaac 480
gggtcttgag gggtttttg ctgaaaggag gaactatatc cggataacct ggcgt 535

<210> 157
<211> 551
<212> DNA
<213> Homo sapien

<400> 157
cgcaagttcc gcctggactg cccgctggcc atggagcgga tcaaggagga ccggcccatc 60
accatcaagg acgacaaggg caacctcaac cgctgcatcg cagacgtggt ctcgctcttc 120
atcacggtca tggacaagct gcgcctggag atccgcgcca tggatgagat ccagcccgac 180
ctgcgagagc tgatggagac catgcaccgc atgagccacc tcccacccga ctttgagggc 240
cgccagacgg tcagccagtg gtgggtgtcc ctcccagcca ggcagagccc cgcagtgccc 300
gagaccctcc cagccaggcg gagccccgca gtgcccctga ggccctcagc ccccacatgc 360

cctgtgctgc actcccaggc tgcagaccct gagcggcatg tcggcgtcag atgagctgga 420
cgactcacag gtgcgtcaga tgctgttcga cctggagtca gcctacaacg ccttcaaccg 480
cttcctgcat gcctgagccc ggggcactag cccttgcaca gaagggcaga gtctgaggcg 540
atggctcctg g 551

<210> 158
<211> 551
<212> DNA
<213> Homo sapien

<400> 158
agtcacatga tgattatgtt tttgtttaac attctttcca tgcacttgtt attttattaa 60
tttgcctgaa tgatgagacc agaccagtgt ctacagattt tcattgtcag aaaaatctat 120
aagtctgccc ttttttacaat gatgatttaa aaaaaacaac agcgtaaata ttagcccaca 180
agagcagtcc taaacaatca caattacact gtactaccca agaagactgt ttattgtgaa 240
gcatttacct ttcaaaaaat cattacattt ctatttcttg gtggagcagc acattgtgga 300
gtgtgattct taattcttca ttgagtttgt caataggaca ttgatgctgg ataggttgtc 360
ttttgttttt atgtctcaga ccatcttgtg agattgtttg cctatctcat aatacagttt 420
tatgcagaaa ggttgaaact atgtaaatgg tttttatgga aattatcagt tacaatattt 480
taaaggtgta gaatggcatc tttgtttata ggagaacatt tgtaaataaa gttaaatttc 540
taagtcaaaa a 551

<210> 159
<211> 541
<212> DNA
<213> Homo sapien

<400> 159
caaggggagc caagtttgcc gatgcagtga atgtggtaaa atattccgga acccaagata 60
cttttctgtg cataagaaaa tccataccgg agagaggccc tatgtgtgtc aagactgtgg 120
gaaaggattt gttcagagct cttccctcac acagcatcag agagttcatt ctggagagag 180
accatttgaa tgtcaggagt gtgggaggac cttcaatgat cgctcagcca tctcccagca 240
cctgaggact cacactggcg ctaagcccta caagtgtcag gactgtggaa aagccttccg 300
ccagagctcc cacctcatca gacatcagag gactcacacc ggggagcgcc catatgcatg 360
caacaaatgt ggaaaggcct tcacccagag ctcacacctt attgggcacc agagaaccca 420
caataggaca aagcgaaaga agaaacagcc tacctcatag ctctcaagcc agttgaagaa 480
accttgcctt ttcagcttga ccctgcaata taacatgcac aggcctgctt gtgaatcagg 540
a 541

<210> 160
<211> 541
<212> DNA
<213> Homo sapien

<400> 160
ttctcttttt cctccagaag tatttgttac aagatttgta aataagagct ctacttagtt 60
tgtttaccat gaacatgttg cagcaaacct tatgcatcta attcctacaa ggttaaagaa 120
aggcttttag acttgccagg ttaagcaaca gccaagttct cagtaattgt ttgccttgat 180
ttatctttta gacttcattt tgccagctct aaaactccca gtcttccttg attttagtcc 240
ttaatctttt atgttctgag caggaagggt aaaagacagg aacctgcttc actgtattaa 300
ctagtccatg ggctgagacc ggggcatctc ttttcttcat actgcaatgt tgctagatac 360
atgatcagac accagagggt tgggcattct tgcaatacct taacagtgct gaaatctgca 420
gcatggtact aaggaagtta aagtttgaat gtaaccactt tatttaaaag gttttttttct 480
ttaatttaaa tgaaatgggg gtgaaagtga acatgatttt gttgaccatg ttcgggaatt 540
a 541

<210> 161
<211> 541
<212> DNA
<213> Homo sapien

<400> 161
agcaggagtt ttgtttgttg ggatttgaac gctacctgca gtgcatcttg ggtgttgaca 60
atatcaaaga tgttatccct ttcccaaggt ttcctcattc atgcctttta tagctggaag 120
attggttaag gaaaagcacc ccccatggca gagacactgc acatgattgt gcatacagca 180
gaatgcatgt ttggatttta gaaatgcaga tttcaatatg taattgttgt gccataagat 240
atcatagaaa aaatataagt ggttgtgatt ttcttagaaa gttgagggta tttcacgtaa 300
ggatgagctc ccgcaagaag aggtacttat agcaagggga ctctcaaatc cattacctca 360
attaagaaat gaagaaattg aattagtctc aaagtttctt ttaaactcta aaacagaatg 420
agataatgta ttttacgttg tctataatca ttaaatcact ccctgtgtaa tttgtgagaa 480
ccatctagta gctcgaaata aaataatgtt gcatcttttc tcccctgcca tatactttgt 540
g 541

<210> 162
<211> 451
<212> DNA
<213> Homo sapien

<400> 162
cgaattctac aagagcctca ctaatgactg ggaagaccac ttggcagtca agcacttttc 60
tgtagaaggt cagttggaat tcagggcatt gctatttatt cctcgtcggg ctccctttga 120
ccttttttgag aacaagaaga aaaagaacaa catcaaactc tatgtccgcc gtgtgttcat 180
catggacagc tgtgatgagt tgataccaga gtatctcaat tttatccgtg gtgtggttga 240
ctctgaggat ctgcccctga acatctcccg agaaatgctc cagcagagca aaatcttgaa 300
agtcattcgc aaaaacattg ttaagaagtg ccttgagctc ttctctgagc tggcagaaga 360
caaggagaat tacaagaaat tctatgaggc attctctaaa aatctcaagc ttggaatcca 420
cgaagactcc actaaccgcc gccgcctgtc t 451

<210> 163
<211> 541
<212> DNA
<213> Homo sapien

<400> 163
cctctggccc ttcgcatacg tgtgtctgct gagtgttcct gcatgtaaga attaagacca 60
agggagggga gagagaaacc cacacataaa caatgcacta aagatcactg aactgtttaa 120
acatttccac ttgccagttt aatttcttga agactgttgc ttgtttggaa tgtttcttgt 180
cactgatttt aaggttgcat ctggaaaaga ctaaaggctt cagtcccctc ccaccaccag 240
aaatgaacaa aaagcatttt acctaaaaat acaccagcaa aatgtactca gcttcaatca 300
caaatacgac tgcttaaaac tgcagaaatt tcctcaacac tcagccttta tcactcagct 360
ggatttttc cttcaacaat cactactcca agcattgggg aacacaactt ttaatcatac 420
tccagtcgtt tcacaatgca ttctaatagc agcgggatca gaacagtact gcatttactt 480
gccaacagaa cagacagacc tgaagtcaag acaactgcat tctctgtgaa gtctggtaaa 540
a 541

<210> 164
<211> 551
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(551)
<223> n = A,T,C or G

<400> 164
cattctttat ttcaggaaag cttgtggctc ttgcagttat tgatgagaaa aatacatcag 60
ttgaacatac cagattgaag tcaattattc aggaagttgc aagagattac agagacctct 120
tccataggta ggtgcttata ttctgagaga ttccttttatt tttgtttaat accattttgt 180
aatttagtaa agcaaactta gtagtattca tgcttttaga ggtgaatcaa agaaatgaag 240
ttaaataaca gtaatcctcg cttatccaca ggagatacat tccaagaccc ccagngggtg 300
actgaaactt agcatagaat ggaatgctac atatactata attttgtctg tacctacata 360
cttatgataa agtttaattt atacattagg tacagtaaga gatcaacaac agtaattaat 420
aataaaatag aacaattata acaatacagt atagtaaaaa ttatgtatat gtttatttt 480
ggaatttttc atttaatatt ttcagaccat ggttgaccat ggggaactga aaccacagaa 540
agtgaatcta t 551

<210> 165
<211> 551
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(551)
<223> n = A,T,C or G

<400> 165
gcgtcacgtg gcccacggcg tccagggcga ccagccgcgg gccgcagggc atggaccttc 60
aggccgccgg ggcccaggcg cagggggccg cggagccgtc tcggggcccg ccgctgccta 120
gcgcgcgggg ggcgcccccc agcccggagg ctggctttgc tacagctgac cactccggtc 180
aggagagaga gactgagaag gctatggatc gactagcccg tggaacacag agcattccta 240
atgacagtcc tgcccggggt gagggcaccc attctgaaga ggaaggcttt gccatggatg 300
aggaggactc tgatggagaa ctgaatacct gggagctgtc agaagggaca aactgtccac 360
ccaaggaaca gcctggcgat ctttttaatg aggactggga ctcggagttg aaagcagatc 420
aagggaatcc atatgatgct gacgacatcc aggagagcat ttctcaagag cttaaacctt 480
gggtgtgctg tgccccacaa ggagacatga tctatgaccc cagctggcac catccgcctn 540
cactgatacc c 551

<210> 166
<211> 536
<212> DNA
<213> Homo sapien

<400> 166
gtctaaaagc tggtgttatt gctgttattg tggttgtggt gatagcagtt gttgctggaa 60
ttgttgtgct ggttatttcc agaaagaaga gaatggcaaa gtatgagaag gctgagataa 120
aggagatggg tgagatgcat agggaactca atgcataact atataatttg aagattatag 180
aagaagggaa atagcaaatg gacacaaatt acaaatgtgt gtgcgtggga cgaagacatc 240
tttgaaggtc atgagtttgt tagtttaaca tcatatattt gtaatagtga aacctgtact 300
caaaatataa gcagcttgaa actggcttta ccaatcttga aatttgacca caagtgtctt 360
atatatgcag atctaatgta aaatccagaa cttggactcc atcgttaaaa ttatttatgt 420
gtaacattca aatgtgtgca ttaaatatgc ttccacagtg cttgcggccg cactcgagcc 480
cgggtgaatg attgagttta aaccgctgag caataactag cataaccccct tggggc 536

<210> 167
<211> 541
<212> DNA
<213> Homo sapien

<400> 167
gcgcgcctcg tccccgcccg accgcatcga catcttcggg cgcacggtga gcaagcgcag 60
cagcctggac gagaagcaga agcgagagga ggaggagaag aaagcggagt tcgagcggca 120
gcgaaaaatt cgacagcaag aaatagaaga aaaactcatc gaggaagaaa cagcacgaag 180
agtagaagaa ttggtagcaa aaagggtgga ggaagaactg gagaaaagga aggatgaaat 240
tgaacgagaa gttctccgaa gggtggagga agccaaacgc atcatggaaa agcagttgct 300
cgaagaactc gagcgacaga gacaagctga gcttgccgca caaaaagcta gagaggtaac 360
gctcggtcgt ttggaaagta gagacagtcc atggcaaaac tttcagtgtt gggttctgcc 420
tcctgctcag ttcagaaaga gatggaatac agactatcta attcctttct cgtctaaact 480
taacattgct gcgaaagtta atttttagc ctattcagaa gtgctgactg ataacttaaa 540
a 541

<210> 168
<211> 551
<212> DNA
<213> Homo sapien

<400> 168
atagacttcc aatcagaagt ctcactggtg gggctggggg tgggggcagg caggaggcat 60
ggatgggaac ctgagtaggt agtgtggcca agagatcagc acaacctttg caggctgact 120
tgctaagtct gacagtgaca aacttgtgag cttactgcag tcagtcacag aggctgttct 180
ttttcacaca ccccttcatg cccggctttc cccatatcca catgcagagg gcgagctcat 240
aaaactacag ggaagcgtga aatgatggct ttggtagctg tttactgggt aaccccactg 300
tgacactgtc cttttcatgt gatgtggaaa cctacttctg tcctccaaac catgaaatgt 360
gtcatctaga ctgcagagta cttgagtgct ttgcctcccg atatgccaga gcttgtggtc 420
caaagcccat tcctgtgtgt ccgtcctgcc atttagccac agaaggctgc ggagtgaggc 480
ggcagctagc ctggccagtg gctgtcccgt ggaccgacac ctgcgccccc ttctgcaagc 540
aggattttct g 551

<210> 169
<211> 551
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(551)
<223> n = A,T,C or G

<400> 169
tgtgctctcc ctcctttcct ctnccttgag cttggttctg cccagcactc gtgcttgttc 60
acataattag gtttcccacc ccagcctacc cgacttactt gctagtctct atgaggtcct 120
tattgcactt attggggttg aagctcttca gaggagctgg aactgtctac cccagggaca 180
cacccatttc gttgctaccc aagtggattc tgagacaggc accatctcct tgttcccct 240
ctctcttttg cctcccactg actgcccttt tccatgtgtc ttcattctgc ctgaagaagg 300
ctttcccagg atgcacgtcc tcagagggag cagcctatct cccccaagct ggaggcggca 360
gaggactggg ccaagcccca acctgcctcc cagccaggct cctccaggcc tctggtttag 420
cggagccccc tgagcccagg cctgtgtcta gccccagtgg ctcactgaac tttcagggca 480
gtcaggggt cctgcttaga agccagtcac cagccctctg cctgcagcca tggaaggggg 540
tgtgcacgtg c 551

<210> 170
<211> 551
<212> DNA
<213> Homo sapien

<400> 170
gttaagagga tccgcgagga gagtggcgcg cggatcaaca tctcggaggg gaattgtccg 60
gagagaatca tcactctgac cggccccacc aatgccatct ttaaggcttt cgctatgatc 120
atcgacaagc tggaggaaga tatcaacagc tccatgacca acagtaccgc ggccagcagg 180
cccccggtca ccctgaggct ggtggtgccg gccacccagt gcggctccct gattgggaaa 240
ggcgggtgta agatcaaaga gatccgcgag agtacggggg cgcaggtcca ggtggcgggg 300
gatatgctgc ccaactccac cgagcgggcc atcaccatcg ctggcgtgcc gcagtctgtc 360
accgagtgtg tcaagcagat ttgcctggtc atgctggaga cgctctccca gtctccgcaa 420
gggagagtca tgaccattcc gtaccagccc atgccggcca gctccccagt catctgcgcg 480
ggcggccaag atcggtgcag cgacgctgcg ggctaccccc atgccaccca tgacctggag 540
ggaccacctc t 551

<210> 171
<211> 551
<212> DNA
<213> Homo sapien

<400> 171
atgcagctca gttctgcaca ggtggagcag ctgcgccagg ccattgaaga actgtactac 60
tttgaatttg tggtagatga cttgccaatc cggggctttg tgggctacat ggaggagagt 120
ggtttcctgc cacacagcca caagatagga ctctggaccc atttggactt ccacctagaa 180
ttccatggag accgaattat atttgccaat gtttcagtgc gggacgtcaa gccccacagc 240
ttggatgggt tacgacctga cgagttccta ggccttaccc acacttatag cgtgcgctgg 300
tctgagactt cagtggagca tcggagtgac aggcgccgtg gtgacgatgg tggtttcttt 360
cctcgaacac tggaaatcca ttggttgtcc atcatcaact ccatggtgct tgtgttttta 420
ctggtgggtt ttgtggctgt cattctaatg cgtgtgcttc ggaatgacct ggctcggtac 480
aacttagatg aggagaccac ctctgcaggt tctggtgatg actttgacca gggtgacaat 540
ggctggaaaa t 551

<210> 172
<211> 541
<212> DNA
<213> Homo sapien

<400> 172
aggatgctgc aagataggaa attctcatag aaattagaaa cctagtcaga ggacaagctt 60
catacagtat gtacagttgg aactgttcaa gtatagtttc agtgtaaaaa gtgctacaat 120
aacaaaccac atttaagaaa gagttcttag tagagaaaca ataagacaaa ataccaaaca 180
tagtacacaa caaatttatg cctcagctac atgatctaaa agttaaaggt cccaggagcc 240
ccatcctgaa cttggaaagt gtagccttca gaggtagttt ctggcacaac gttttgatct 300
tcctcttcct ggaaatatat taaaaaataa gtataaaaat aataagtatt ccaagcagca 360
gcctacctga agtctgtatt taatctctag gtcttgatct gcatattaca cctttaatcc 420
tgtatggtat tgtagtatct gacatagggg agggagggtc tctgaaactt tgccaattca 480
taagtgctag ctactacagt aacagaaaca ggccgagatt ttttttcttc cccaaccgtc 540
a 541

<210> 173
<211> 522
<212> DNA

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(522)
<223> n = A,T,C or G

<400> 173
agctctnctg ccctccgctg tcactcctcg gaaagccaaa ttaggtgaca ctaaagagct 60
cgaagacttc attgccgatc tggacagaac cttagcaagt atgtgaagca aggagtttgg 120
ggtccagaag gctccgagga cctggcaaat cggctactag aatctgctgt gggagagggc 180
agagctgagg ctcctgcccc ctggccattc ttggttcact ataacattag ccattgggcc 240
catctctggg cagttcggag agtgaagctc actttgttta cctacctgca gcatattcaa 300
cagaggatct atctaatgtg tttttactcc tttaaacata gcccttctat aatttaaaat 360
gcttttatgg aaatatttgt aattacttat atatagttgg aggtcataat aagctttccc 420
atcatagtat atttttgtat gcaaataaaa ttaaaacgga gatctgtaaa aaaagcttgc 480
ggccgcactc gagcccgggt gaatgattga gtttaaaccg ct 522

<210> 174
<211> 427
<212> DNA
<213> Homo sapien

<400> 174
attattctaa ataaaaggaa aaaggcttac actacctaaa gctgtgctct ctgcctcctg 60
ggagagggcc gcaaagccag gcaccccgcc aaccactggg ggtcctaatc cacctgctgg 120
gcatcacctc tcctcctcct cagaattggg tgtttgctga ccatcaaaag caatgacttt 180
ttattctgtt tgtactgaac caaaacaaac aactgtgtat agactgctgt tttcttttt 240
atttgaaatg aggcattttg gtgttctttc ccctaccata cggcctgtct gcccttccct 300
ccccacattg gctccagcag agtagccgaa ggtcctgccg ccgccgccac caccaccacc 360
actgcagcaa caacagcagc agcagcagca gcgcctgcat agctccactc tgacctgtga 420
aggaatg 427

<210> 175
<211> 451
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(451)
<223> n = A,T,C or G

<400> 175
agccgctcca gccatgctct ccgnnctcgc ccggcctgcc agcgctgctc tccgccgcag 60
cttcagcacc tcggcccaga acaatgctaa agtagctgtg ctaggggcct ctggaggcat 120
cgggcagcca ctttcacttc tcctgaagaa cagccccttg gtgagccgcc tgaccctcta 180
tgatatcgcg cacacacccg gagtggccgc agatctgagc cacatcgaga ccaaagccgc 240
tgtgaaaggc tacctcggac ctgaacagct gcctgactgc ctgaaaggtt gtgatgtggt 300
agttattccg gctggagtcc ccagaaagcc aggcatgacc cgggacgacc tgttcaacac 360
caatgccacg attgtggcca ccctgaccgc tgcctgtgcc cagcactgcc cggaagccat 420
gatctgcgtc attgccaatc cggttaattc c 451

<210> 176
<211> 540

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(540)
<223> n = A,T,C or G

<400> 176
agagaccagg acccaggact ttccncttcc agtccacagc ctttcatttt agccatggtt 60
ctgggcctag ctgattcaga ttcagtgggc tggggtagat ggagcctggg tgtctgtacc 120
ttttgtaagt ttcgcagtaa attcagatta tagccaggtt tagcacttgc tgtgaggaat 180
gtactgcctc tgtgtaagcg gcagtggaat gtgggaagcc agatcggatt ctggagaatg 240
atgacttgac cagagcagaa agagggtcat gaacacaggt gatcaaaggt ggtcgtttgt 300
tcatgtgggc ctgacaggga gctgctggca ggtctaggtg tgacttggag gccgctgggt 360
acctaagcct ctatttcctc cttgctgagc tttgggagca ccgtgggctg caacttcctc 420
cctggcagat cccagtagat ctgtgcgcag ccaagtgagt gggaaggcac tcatacagca 480
ggcccagggc cagcaccacc tgaccattcc tgcttacctg tcagcgctct gttctcagat 540

<210> 177
<211> 451
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(451)
<223> n = A,T,C or G

<400> 177
aagcgcggca tggaggaggc ggatnccgcg gcgagccggg ccgagcagtg agggccctag 60
cggggcccga gcggggcccg gggcccctaa gccattcctg aagtcatggg ctggccagga 120
cattggtgac ccgccaatcc ggtatggacg actggaagcc cagccccctc atcaagccct 180
ttggggctcg gaagaagcgg agctggtacc ttacctggaa gtataaactg acaaaccagc 240
gggccctgcg gagattctgt cagacagggg ccgtgctttt cctgctggtg actgtcattg 300
tcaatatcaa gttgatcctg gacactcggc gagccatcag tgaagccaat gaagacccag 360
agccagagca agactatgat gaggccctag gccgcctgga gcccccacgg cgcagaggca 420
gtggtccccg gcgggtcctg gacgtagagg t 451

<210> 178
<211> 643
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(643)
<223> n = A,T,C or G

<400> 178
agcgcaatgt ggggctccca ggtcccattn gtcccctggg ggacagcggc ccccaaggac 60
ngaagggcgt gaaaggcaat ccaggcaata tcagggacca ncccggccca gctttctcag 120
ccattcggca gaacccaatg acgcttggca acgtggttat ctttgacaag gtcctcacca 180
accagganag tccataccag aaccacacgg gtcgcttcat ctgtgcagtg cccggcttct 240
attacttcaa cttccaaagt gatctccaag tgggaccttt gtctgtttat caagtcttcc 300

tccgggggcc agcccaggga ttccctgagt ttctctaaca ccaacaacaa ggggctcttc 360
caggtgttag caggggggcac ccgtgcttca nctgcgacca aggggaccaa ggtgtggatc 420
gagaaggacc ccgcaaaggg tcgcatttac cagggcactg aacccgacag cattttttcag 480
cggattcctc attttccccct cggcctgagc tggggatctg cccctgcatc ctgccatctc 540
ctgcgctccc tgttgtggac cacgcccccc ntccgcctga ccctccctcc gaattttgca 600
aatgaagggg ctggggcttt aacaccctgg gggaggggggt tcg 643

<210> 179
<211> 651
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(651)
<223> n = A,T,C or G

<400> 179
gtatggttta ttctgaggtt cctaacttna gtgagccaaa cccagagtac agcacccagc 60
nggcacccaa caaggcggtg cagaacgaca gcaacccttc agcttcccag cctaccactg 120
gaccctctgc tgcctctcca gcctctgaga accagaatgg gaatggactg agtgccccac 180
caggtcccgg tggtggccca catccccctc atactccctc ccacccaccc agcacccgaa 240
tcactcgaag ccagcccaac cacacacctg caggcccgcc tggcccttcc agcaacccctg 300
ttagtaacgg caaagaaacc cggaggagca gcaagagata gcatgacatt ctttcttcct 360
gccaccaacc acatcccaag tgtcccctgg agagcaagat agccttccac tgattggctg 420
gtgtagcagt attttagcca ctgaacttca gtggagggtg gtgagcagtg tccttatcca 480
ccctaatctc atactccctc attgtccagc tgaactacct gtcccctggg agtcaggacc 540
ctctgcctgc tctctttcct ctttagaaat ggcagttact ggctgggcgc agtggctcac 600
gcttgtaatc ccagcacttt gggaagccga ngtgggcgga tcacctgagg t 651

<210> 180
<211> 651
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(651)
<223> n = A,T,C or G

<400> 180
agcctaccag aacatgcggt cctcaaantg aaaggaaact ttaccctccc agaggtagct 60
gagtgctttg atgaaataac ctatgttgaa cttcagaagg aagaagccca aaaactcttg 120
gagcaatata aggaagaaag caaaaaggct cttccaccag aaaagaaaca gaacactggc 180
tcaaagaaaa gcaataaaaa taagagtggc aagaaccagt ttaacagagg tggtggccat 240
agaggacgtg gaggattcaa tatgcgtggt ggaaatttca gaggaggagc ccctgggaat 300
cgtggcggat ataataggag gggcaacatg ccacagagag gtggtggcgg tggaggaagt 360
ggtggaatcg gctatccata ccctcgtgcc cctgtttttc ctggccgtgg tagttactca 420
aacagaggga actacaacag aggtggaatg cccaacagag ggaactacaa ccagaacttc 480
agaggacgag gaaacaatcg tggctacaaa aatcaatctc agggctacaa ccagtggcag 540
cagggtcaat tctggggtca gaagccatgg agtcagcatt atcaccaagg atattattga 600
atacccaaat aaaacgaact gatacatatt tctccaaaac cttcacaaga a 651

<210> 181
<211> 631

<212> DNA
<213> Homo sapien

<400> 181
agcggaagga gcccggccgc ccccgccggc tgactaccac cgggccatca cgcttctctc 60
caggcactct tccaagtgga gtacgcgcag gaggccgtca agaagggctc gaccgcggtt 120
ggtgttcgag gaagagacat tgttgttctt ggtgtggaga agaagtcagt ggccaaactg 180
caggatgaaa gaacagtgcg gaagatctgt gctttggatg acaacgtctg catggccttt 240
gcaggcctca ccgccgatgc aaggatagtc atcaacaggg cccgggtgga gtgccagagc 300
caccggctga ctgtggagga cccggtcact gtggagtaca tcacccgcta catcgccagt 360
ctgaagcagc gttatacgca gagcaatggg cgcaggccgt ttggcatctc tgccctcatc 420
gtgggtttcg actttgatgg cactcctagg ctctatcaga ctgacccctc gggcacatac 480
catgcctgga aggccaatgc cataggccgg ggtgccaagt cagtgcgtga gttcctggag 540
aagaactata ctgacgaagc cattgaaaca gatgatctga ccattaagct ggtgatcaag 600
gcactcctgg aagtggttca gttcaggtgg c 631

<210> 182
<211> 559
<212> DNA
<213> Homo sapien

<400> 182
caacatacct caacttctgc cgctccctgc ggtttgacga caagcccgac tactcttacc 60
tacgtcagct cttccgcaac ctcttccacc ggcagggctt ctcctatgac tacgtctttg 120
actggaacat gccgaaattc ggtgcagccc ggaatcccga ggatgtggac cgggagcggc 180
gagaacacga acgcgaggag aggatggggc agctacgggg gtccgcgacc cgagccctgc 240
cccctggccc acccacgggg gccactgcca accggctccg cagtgccgcc gagcccgtgg 300
cttccacgcc agcctcccgc atccagccgg ctggcaatac ttctcccaga gcgatctcgc 360
gggtcgaccg ggagaggaag gtgagtatga ggctgcacag gggtgcgccc gccaacgtct 420
cctcctcaga cctcactggg cggcaagagg tctcccggat cccagcctca cagacaagtg 480
tgccatttga ccatctcggg aagtgaggag agcccccatt ggaccagtgt ttgcttagtg 540
tcttcactgt attttcttt 559

<210> 183
<211> 651
<212> DNA
<213> Homo sapien

<400> 183
acaagacatc ctccccctcc agtacggaag ttccaaggca cttgttttcc agcatatcag 60
cctaacctca gtgccttgaa atatggcttt aagcctttga gaactgagat ttcctgaaac 120
cataggccct tgccccaggg gtttctccac atccgggtgt taagacacct gatggcactg 180
ttggtttgtc ccctataccc cagaaaatct atcctgcaag gtagctactt caatcttgtc 240
attaaaatgt gtcaagtcac agctcgcaat gccaaaggaa tgctggggca gtaagtgagg 300
tggataagtg agaagggcct ggtggtgaaa gcggccaggg accagaatgc tccagaccta 360
cagagctggt caaggttaag tgccttaaac ttaccaatcc tgggctcagt tctcctttga 420
aaggagaaag tctttgtcct ctacttaggc agctgggcta gaagtgcctt ttgacttcta 480
atgttaacta ccctccaaag cctcctgggt caagaaggct ctcccaaact ccacccctgt 540
tcttcctggt cagagaacca gtcagtcatt ctagtcttct agtccttaaa ctgatctgat 600
gacttggaac ataggatttc actgcaagtc ttggcttttt agtctgggaa a 651

<210> 184
<211> 577
<212> DNA
<213> Homo sapien

<400> 184
agccacacac cacctgtcca aaaaggcctt cgatacggga taatcctatt tattacctca 60
gaagtttttt tcttcgcagg atttttctga gccttttacc actccagcct agcccctacc 120
ccccaattag gagggcactg gcccccaaca ggcatcaccc cgctaaatcc cctagaagtc 180
ccactcctaa acacatccgt attactcgca tcaggagtat caatcacctg agctcaccat 240
agtctaatag aaaacaaccg aaaccaaata attcaagcac tgcttattac aattttactg 300
ggtctctatt ttaccctcct acaagcctca gagtacttcg agtctccctt caccatttcc 360
gacggcatct acggctcaac atttttttgta gccacaggct tccacggact tcacgtcatt 420
attggctcaa ctttcctcac tatctgcttc atccgccaac taatatttca ctttacatcc 480
aaacatcact ttggcttcga agccgccgcc tgatactggc attttgtaga tgtggtttga 540
ctatttctgt atgtctccat ctattgatga gggtctt 577

<210> 185
<211> 340
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(340)
<223> n = A,T,C or G

<400> 185
caaattattt ctgaatgtca gtttatgcca gagtcatgct cactttgcag ggtggagcac 60
ttctcgcgcc taatggttgg ttggtggtct tggcctctag tttcactctt tgaggataga 120
gtgccatctg cagccttcac tgttgggatc tgagctgtag tggctgtgaa tagcctagag 180
tggtcaggca gcccacctcg gcactcctgg ggtgggggtg gggacagctg cttaaccttt 240
attcttggtc ggtctgtcgg caactttggg naccaccagt aggatgtggt taagattcag 300
ttcttgctga gctaaggaag catttctcat ttctttttaa 340

<210> 186
<211> 541
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(541)
<223> n = A,T,C or G

<400> 186
gaaaaggcaa attacgtgac agcagatcan cgcttgcctg gggctgggcc tgaggcggga 60
aagatggcaa agggatgcaa gggcccccta gagtgaggaa aacaccctaa aaggctgctc 120
tgcaaccacg tggtgaatac ttactaggat tcaatgaaga gtacacttaa agtgagtgaa 180
ttttatggta ggtaacttat acctaaataa agctattttt aaagtatagt ggataaatat 240
gcatttatat ttttaattga ctatttttatc tatctttttt ctatttttttt tttaaaggct 300
ggtgcctgga ctgttagaaa tgcaaaatgc ttgttcccca gtgccacaga gaaatagcac 360
tcgagcatcc atttaatgtt ctcagcaagg cagttttgac tttttgcata aagggtgacc 420
ctcacaggtg gagccatggc aagagcatac ctggacacgg gagggacagg aggtcttatc 480
cctgaggcag gtagtcccta ctgctgtgtc gttcccttat tggctagggt tggaccacac 540
a 541

<210> 187
<211> 1459

<212> DNA
<213> Homo sapien

<400> 187

tggagcagcc ccaccacaag aaggagtgct acctgaactt cgatgacaca gtgttctgcg 60
acagcgtatt ggccaccaac gtgacccagc aggagtgctg ctgctctctg ggggccggct 120
ggggcgacca ctgcgaaatc tacccctgcc cagtctacag ctcagccgag ttccacagcc 180
tctgcccaga cggaaagggc tacacccagg acaacaacat cgtcaactac ggcatcccag 240
cccaccgtga catcgacgag tgcatgttgt tcgggtcgga gatttgcaag gagggcaagt 300
gcgtgaacac gcagcctggc tacgagtgct actgcaagca gggcttctac tacgacggga 360
acctgctgga atgcgtggac gtggacgagt gcctggacga gtccaactgc cggaacggag 420
tgtgtgagaa cacgcgcggc ggctaccgct gtgcctgcac gcccctgcc gagtacagtc 480
ccgcgcagcg ccagtgcctg agcccggaag agatggacgt ggacgagtgc caggacccgg 540
cagcctgccg ccctggccgc tgcgtcaacc tgccgggctc ctaccgctgc gagtgtcgcc 600
cgccctgggt gcccgggccc tccggccgcg attgccagct ccccgagagc ccggccgagc 660
gtgccccgga gcggcgcgac gtgtgctgga gccagcgcgg agaggacggc atgtgcgctg 720
gcccccctggc cgggcctgcc ctcaccttcg acgactgctg ctgccgccag ggccgcggct 780
ggggcgccca atgccgaccg tgcccgccgc gcggcgcggg gtcccattgc ccgacatcgc 840
agagcgagag caattccttc tgggacacaa gccccctgct gttggggaag cccccaagag 900
atgaggacag ttcagaggag gattcagacg agtgtcgctg cgtgagtggc cgctgcgtgc 960
cgcggccggg cggcgccgtg tgcgagtgtc ccggcggctt ccagctcgac gcctcccgcg 1020
cccgctgcgt ggatatcgac gagtgccgag agctgaacca gcgcgggctg ctgtgcaaga 1080
gcgagcgctg cgtgaacacc agcggctcct tccgctgcgt ctgcaaagcc ggcttcgcgc 1140
gcagccgccc gcacggggcc tgcgttcccc agcgccgccg ctgacgccgc cgacgccgcc 1200
ctcggcccag acctcggtga tcactgaggg atttccgcga gctcggcctc actctgccc 1260
cgacttgtgg ctcggaccca gggaccttca gggcccgcag accctcccgg cgccttgaga 1320
cccgaggcgc ccctaccggc ccccctcccc ggttagcggg cggttgtaag gtctccggcg 1380
ggcgctgcct gccttcctcc cagagggtgt ttcctagaaa ctgataaatc agatcgtgcc 1440
tctttaaaaa aaaaagctt 1459

<210> 188
<211> 2514
<212> DNA
<213> Homo sapien

<400> 188

gccgcgcgtc ttgccctacc accctgcaca gcccggccag gccgccaaaa aggccgtcag 60
gacccgctac atcagcacgg agctgggcat caggcagagg ctgctggtgg cggtgctgac 120
ctctcagacc acgctgccca cgctgggcgt ggccgtgaac cgcacgctgg ggcaccggct 180
ggagcgtgtg gtgttcctga cgggcgcacg gggccgccgg gccccacctg gcatggcagt 240
ggtgacgctg ggcgaggagc gacccattgg acacctgcac ctggcgctgc gccacctgct 300
ggagcagcac ggcgacgact ttgactggtt cttcctggtg cctgacacca cctacaccga 360
ggcgcacggc ctggcacgcc taactggcca cctcagcctg gcctccgccg cccacctgta 420
cctgggccgg ccccaggact tcatcggcgg agagcccacc cccggccgct actgccacgg 480
aggctttggg gtgctgctgt cgcgcatgct gctgcaacaa ctgcgccccc acctggaagg 540
ctgccgcaac gacatcgtca gtgcgcgccc tgacgagtgg ctgggtcgct gcattctcga 600
tgccaccggg gtgggctgca ctggtgacca cgaggggtg cactatagcc atctggagct 660
gagccctggg gagccagtgc aggaggggga ccctcatttc cgaagtgccc tgacagccca 720
ccctgtgcgt gaccctgtgc acatgtacca gctgcacaaa gctttcgccc gagctgaact 780
ggaacgcacg taccaggaga tccaggagtt acagtgggag atccagaata ccagccatct 840
ggccgttgat ggggaccggg cagctgcttg gcccgtgggt attccagcac catcccgccc 900
ggcctcccgc tttgaggtgc tgcgctggga ctacttcacg gagcagcatg ctttctcctg 960
cgccgatggc tcaccccgct gcccactgcg tggggctgac cgggctgatg tggccgatgt 1020
tctggggaca gctctagagg agctgaaccg ccgctaccac ccggccttgc ggctccagaa 1080
gcagcagctg gtgaatggct accgacgctt tgatccggcc cggggtatgg aatacacgct 1140

ggacttgcag ctggaggcac tgacccccca gggaggccgc cggcccctca ctcgccgagt 1200
gcagctgctc cggccgctga gccgcgtgga gatcttgcct gtgccctatg tcactgaggc 1260
ctcacgtctc actgtgctgc tgcctctagc tgcggctgag cgtgacctgg cccctggctt 1320
cttggaggcc tttgccactg cagcactgga gcctggtgat gctgcggcag ccctgaccct 1380
gctgctactg tatgagccgc gccaggccca gcgcgtggcc catgcagatg tcttcgcacc 1440
tgtcaaggcc catgtggcag agctggagcg gcgtttcccc ggtgcccggg tgccatggct 1500
cagtgtgcag acagccgcac cctcaccact acgcctcatg gatctactct ccaagaagca 1560
cccgctggac acactgttcc tgctggccgg gccagacacg gtgctcacgc ctgacttcct 1620
gaaccgctgc cgcatgcatg ccatctccgg ctggcaggcc ttctttccca tgcatttcca 1680
agccttccac ccagctgtgg ccccaccaca agggcctggg cccccagagc tgggccgtga 1740
cactggccgc tttgatcgcc aggcagccag cgaggcctgc ttctacaact ccgactacgt 1800
ggcagcccgt gggcgcctgg cggcagcctc agaacaagaa gaggagctgc tggagagcct 1860
ggatgtgtac gagctgttcc tccacttctc cagtctgcat gtgctgcggg cggtggagcc 1920
ggcgctgctg cagcgctacc gggcccagac gtgcagcgcg aggctcagtg aggacctgta 1980
ccaccgctgc ctccagagcg tgcttgaggg cctcggctcc cgaacccagc tggccatgct 2040
actctttgaa caggagcagg gcaacagcac ctgaccccac cctgtccccg tgggccgtgg 2100
catggccaca ccccacccca cttctccccc aaaaccagag ccacctgcca gcctcgctgg 2160
gcagggctgg ccgtagccag accccaagct ggcccactgg tcccctctct ggctctgtgg 2220
gtccctgggc tctggacaag cactgggggа cgtgccccca gagccaccca cttctcatcc 2280
caaacccagt ttccctgccc cctgacgctg ctgattcggg ctgtggcctc cacgtattta 2340
tgcagtacag tctgcctgac gccagccctg cctctgggcc ctgggggctg ggctgtagaa 2400
gagttgttgg ggaaggaggg agctgaggag ggggcatctc ccaacttctc ccttttggac 2460
cctgccgaag ctccctgcct ttaataaact ggccaagtgt gaaaaaaaaa gctt 2514

<210> 189
<211> 2658
<212> DNA
<213> Homo sapien

<400> 189
gggacacgcg cgaggagccg ccgcagccgc cgccgccgcg ctgtggagcc cgagtgagcg 60
cggcgccagc ccgtccggcc gctggacaac atggaggctg cactgcccgg gccgccgtgg 120
ccgctgctgc tggcgctgtg cggctgcccg gcccccgccg caggtacatg tactgggcag 180
attggggtga gacgccctgg atcgagcagg tggggatgga cgacagcacc cagaagatca 240
ttgtggactt ggacatttac tggcccaata gactgaccat cgacctgggg gagaagaagc 300
tctactgggc tgacactaag ctcggcttca tccaccgtgc caacctggac ggctcgttcc 360
ggcagaaggt ggtgaagggc agcctgacgc acccttcac cctgacactc tccggggaca 420
ctctgcactg gatggactgg ctgacgtgct ccatccatgc ctgcaacagg aacactgcgg 480
ggaagaggaa cgagatcttg agcgccctct actcagcatg gatatccagg tgctgagcca 540
ggagcagcag cctttctgag ccaaggaggt cctgctgctg gcccggtaga cggaccaatg 600
gagggtctcg ctggacacac cggacttcac cgacatcgtg ctgcagctgg acgacatcca 660
gcatgccatc gccatcaact atgacccaat ggagggccac gtctactgga cggatgacga 720
ggtgtgggcc atccgcaggg catacctgga tgggtctggg gcgcagacgc tgatcaacac 780
caagatcaac gaccccgatg acatcgcagt caactgggtg gcccgaagcc tctattggac 840
ccacacaggc actgaacaca tcgaggtgac gtgcctcaac agcacctccc acaagatcct 900
ggtgtcagag gacatggacg agccccgtgc cattgcactg catcccgaga tgggcctcac 960
gtactggata gactggggag agaaccctga aatcaagcgt gccaacctgg acaggcagga 1020
gctgcgcgtg ctggtcaatg cttccctcgg gtggcccaac ggcctggccc tggacctgca 1080
ggaagggaag ctctactggg gagatgccaa gacagacaag atcgaggcga tcagtgtcga 1140
tgagacgaag aggcagaccc tcctgaagga caagctccca cacattttca ggttcaccct 1200
gctgggggac ttcatctact ggaccgcctg gcagcaccac agcatcaagc gggtacacaa 1260
ggtcaaggcc aaccgggacg tcatcattga ccagctgccc gacctgatgg ggctcaaagc 1320
tgtgaacgtg gacaaggtcg tcggtgagtc caggggggtg ggccccaagc cgttgctcag 1380
ctgcagactt gcatgaggaa gaagtgacag ttccaaacct gggcataagt gttgagctgg 1440
ggtgccctgc cctggggaag ggcaggacag gaaaggtgac agtatctggc taaggataga 1500

tgggaaggga ccaagggtgc tgattaggga atggttatgg actaggagta tcagtaacaa 1560
tggttagaaa gtggctaaca tttgttgagc acctgctgtg tgcctggccc cggctgggag 1620
ccttcgtgcc cagagtgacc ccgtctgaat gcagttcttt gcctcattca aactggggag 1680
tgggaggcag agctgcacag ctcacgggtg ccgagctcag gactcactcc tgggtctgcc 1740
tgggctgggc tgtgcttgtt gcccctgtgg ccaacacacg cgcacctttc acctgaaagc 1800
caggatccgc agaacgttcc ccgaggaggt cgttgtttgg cactatgatt tgtctcttcc 1860
taaaaaggtg atagagttac actggagaga gcagcatcca ggtgcagcag ggatgggcct 1920
ggggctcacg ggcaggggct ctgtgtccgg ctggggcctg gggtcctgcg ctgcacctgt 1980
gtgtcagaag cactcagtaa atctttgctg atgaaggatg agaggataga ggacatgatg 2040
cttgctgctg cattgcctgc aatcctggat gagatgccca ggttggcttt gctccccgtc 2100
gggtggatgt gatgttagct gtgatgttag gtccctggct ttaaaatacg agggagctgg 2160
gaattgaggg agcaggttgg ggcagaaagg acagccccgc agaggcctgg agctgaggca 2220
gtgcgggcga ccccaggagc agtgagtgct tccgtcacag ccttcatcgc accctgtggt 2280
cctcataaag gggatggaat ctacgaattt agttttccca gcctccttaa aaactcattc 2340
atgccagggg cagtggctca cacctgaaat cccaccactt tgggaggctg aggcaggctg 2400
attacttgag gtcaggagtt tgagaccagc ctagccaaca tggtgaaacc ccgtgtctac 2460
tcaaagtaca aaaaaaaaaa ttagtcagac gtggtgtcac gcacctgtaa tcccagctct 2520
ttgggaggct gaggcaggag aatcacttga acccaggagg cagaggttat agtgagccag 2580
tattgcgcca ctgacctcca tctgggcaat agagtgagac cctgtctcaa aaaaaaaaaa 2640
aaaaaaaaaa aaaagctt 2658

<210> 190
<211> 2082
<212> DNA
<213> Homo sapien

<400> 190
gggacccgtc cgacctcaag gaggcggtca catacatccg cttccgacac ccggcggcgc 60
cgctgttcgc ggtgagcgaa ggctcgggct cggcgctgct cctgtcctac ctgggcgagt 120
gcggctcctc cagctacgtg acaggcgccg cctgcatctc gcccgtgctg cgctgccgag 180
agtggttcga ggccggcctg ccctggccct acgagcgggg ctttctgctc caccagaaga 240
tcgccctcag caggtatgcc acagccctgg aggacactgt ggacaccagc agactgttca 300
ggagccgttc ccttcgagag tttgaggagg ctctcttctg ccacaccaaa agcttcccca 360
tcagctggga tacctactgg gaccgcaacg acccgctccg ggatgtcgat gaggcagccg 420
tgcctgtgct gtgtatctgc agtgctgacg accccgtgtg tggaccccca gaccacactc 480
tgacaactga actcttccac agcaacccct acttcttcct cctgctcagt cgccacggag 540
gccactgtgg cttcctgcgc caggagccct tgccagcctg gagccatgag gtcatcttgg 600
agtccttccg ggccttgact gagttcttcc gaacggagga gaggattaaa gggctgagca 660
ggcacagagc ttccttcctt gggggccgtc gtcgtggggg agccttgcag aggcgggaag 720
tctcttcctc ttccaacctg gaggagatct ttaactggaa gcgatcatac acaaggtgaa 780
agacctggcc tgagaacccc caagtcctgc aaagaaaaac agagctgggc aaggggagt 840
cctggaaaga tggggcggac tgaacagagg gagctccagc tctgtgctcc tcattcagtc 900
cctctctctt aaattggtgc cttgaaagag aaggaacgtc ctgcgagcct gcactcactt 960
catcctcagc agaactcctg cctggcctct gctcaacata tccctactca tccggtcagc 1020
agcggcgcgt tccagtcact gtcacctgtc actgacatca caagccaaag gatagcactt 1080
tttcaatcca tggactcagg agaaaatgcc ctcttactgg cagtggctag agggatgaga 1140
cgtttgtgta tgtcactggg cagtgacccc gattctcaag ctggagccat ttgatgtcat 1200
gaggacagga tgtttgtgtc tcggccccac ttccctcatt tgctctgtgg ttgtggcgcc 1260
ctgctttgac cgaatgctct ggcaactgcg gcagcaggct tgtgtgtgtg agaagggcgg 1320
cagaggcagt ggggctggct cgctgagggc tgagcaggta gagaatgtga atgtccatct 1380
caaggaatgc gaagttccaa tctctcagag gagctctgca agtctcagct aggaatgggg 1440
ctgatctggc ccccaagcct cactggagag atcccaaggg tattttacaa ataggaagcg 1500
ctctcatgca ggatctgagg gctttgctag accatggttc aaggagtcag aatcttggct 1560
accctacatg ctccctctgg caccatttca ggggattact ccagagagct gcagccagaa 1620
ctgtttatct cctgagatat ggactttctc taagtcagga aaagcacaaa cctcataact 1680

taggtgtttt ctggttggtt gctccctttt ctgatcctac tactggcctc tccctttcca 1740
gagccaggag acacatctgg gaccagagga aacaaaagca ccaagaggtt gactagagct 1800
gaccagtgat ttacaaatat ttcatattcc aatgcttcgt ctgctcaatt attcaaaagg 1860
ccatccctgt aaaggcagaa aagcttcctc ttagatctgg ccactcagca aataagtggg 1920
tatttggttt tatgaagaat aatgctgagg actagagata tagtatatca gagataatat 1980
atagtataca ggttgatcta aatattctgg tttggattag gcaattaaaa tttttttggc 2040
aacaggacat atgtttccat gtgttgctaa gattcaaagc tt 2082

<210> 191
<211> 2004
<212> DNA
<213> Homo sapien

<400> 191
ggaggccagc cccggacccc cagccgcgcc cagggcccgc aggaggctgg gcagcggccc 60
ggacagggag ctccgcaagc cggaggagcc ggagaacggc gagcccacgg ctgcggccac 120
cgccaggagg agcaagaggg agaggcgcga ggaggacagg gccccggcag agcaggtccc 180
gcggagcccg gtcatcaaga tctcctacag cacgccccag ggcaagggag aggtggtcaa 240
gatcccctcc cgcgtgcacg gctctctgga gcccttccgt ccccagcagg ccccgcagga 300
cgacggcagc caggaccccg aggtgctgga cagagagtcc cgggaccggc cgtcctgcgc 360
gccctcggcc tccatcccca agttgaaact gacacggcct gtgccggccg gcgcggacct 420
gccgcccccc aagatccgcc tgaagcccca ccgtctgggg gacagcgagc acgagcccgt 480
gtaccgggcc gagctggtgg gggagctgaa cgggtacctg cgggacagct cgccggcgcc 540
ctgtgcggac ggccctgccg gtgggctggc ggacttgtct tctggaagtt cgggtgagga 600
cgatgacttc aagagctgtc cccagggtcc acagggacgc gagggcttgg cttttctcgt 660
cagctgccct gaggggagag cggactgtgc cagtgagtcg gcgtgcagca gcgacagcct 720
ggacgaggcc agatcgtccg gctcggaagg gacgccggca gacacgggtg acctctcgcc 780
tggccacggc gcgtcagcgc cctcggtgtc cagagaggct cgccaaacgg tgccgccccc 840
gacggtcagg ctgcacacac agagcgtgtc ggagtgcatc acggaggacg gcaggactgt 900
ggccgtgggg gacatcgtct ggggtaagat ccatggtttt ccttggtggc cggcgcgtgt 960
tcttgacatc agtctcggcc agaaggagga cggagagccg tcttggcgag aagcgaaggt 1020
ctcgtggttt ggttctccga ctacgtcgtt cttgtctatt tcaaaactct ccccttttctc 1080
tgaatttttc aaactgagat ttaatcgtaa gaagaagggg atgtatcgga aagctataac 1140
cgaggctgca aatgccgcaa gacacgtggc cccggaaatc agggagctct taacccagtt 1200
tgaaacgtaa ctggttccct gaccaggtca ccgacgatga gggcgcacct gctgtcctgg 1260
agcttccgat ctctgttcgg ggaccctgtg agccttctgg ccggctgcgt gcagagccca 1320
ctgggcacgg tggtcggcct ggtgtgaggc cccccgggga ccggcagtgt gtccagggag 1380
gggatggccc tgagcccagc ggctcctccc gctgagtgta tttcttccca ccactcagcg 1440
catggctgcc ctggctcacg aggcagtcgg gactcgtgac ttgctggctg ctggctgctg 1500
ctgcctcgcg cgctggggtt ccatggagga actgggcctg ccgccccggc ctcacctgct 1560
gcccgcatgc tggggtccca tggaggaacc aggcctggcg ccccggcctc gcccagcggc 1620
tggtgtgggc agctgttccc tgcctcgcag tgtcctggag gcagctgtct cgcagactca 1680
gcttggtctc ccgcaggctt cagaaaaacc caattgcacg tgtgggattc ttccccagcc 1740
cttgggtgcg gggtctctgt gcacttgaga gctgggggac ccacccaccg tctcccactc 1800
caagttcatc ccagagcgtg gggcccgctg ggcacctggc acctggcctg cagatgctgc 1860
tacgagtgac cagtgctgtg tggaggaggc caagggcggc ccctgggaaa ccgggcttca 1920
acagtacaag gaaaagaaac ttgctctgtt ttgtaagagc gtgttccttt tcctcttttt 1980
tgtcacttta aagaccaaaa gctt 2004

<210> 192
<211> 1130
<212> DNA
<213> Homo sapien

<400> 192

```
ctcgtcagga gccaacgtaa gagttcttca ctaccacaac caaggaggga tatgataggc      60
ggccagtgga tataactcct ttagaacaaa ggaaattaac ttttgatacc catgcattgg     120
ttcaggactt ggaaactcat ggatttgaca aaacacaagc agaaacaatt gtatcagcgt     180
taactgcttt atcaaatgtc agcctggata ctatctataa agagatggtc actcaagctc     240
aacaggaaat aacagtacaa cagctaatgg ctcatttgga tgctatcagg aaagacatgg     300
tcatcctaga gaaaagtgaa tttgcaaatc tgagagcaga gaatgagaaa atgaaaattg     360
aattagacca agttaagcaa caactaatgc atgaaaccag tcgaatcaga gcagataata     420
aactggatat caacttagaa aggagcagag taacagatat gtttacagat caagaaaagc     480
aacttatgga aacaactaca gaatttacaa aaaaggatac tcaaaccaaa agtattattt     540
cagagaccag taataaaatt gacgctgaaa ttgcttcctt aaaaacactg atggaatcta     600
acaaacttga gacaattcgt tatcttgcag cttcagtgtt tacttgcctg gcaatagcat     660
tgggatttta tagattctgg aagtagtatt aatgctcatc ctgctgtggc tgttggcttc     720
ttagaacacc aaaccgggag agatttactt tgaacattgt cagttgcagc aaaaatttac     780
tacacaagat tattcgaagt gtatacggac taaaagagga agtgttttag aatgagaaga     840
gatactgtgt ctttattgtg tgtgtgtgag tgcaggtgtg tgtctttatt atattgaaaa     900
gctgtcactc agacctggtt tgagatagaa gagcattttg tccttttgat agttaataga     960
aattgaacca gagttttctt atgtttgctt gaacagttgt gtaaatcata caggattttg    1020
tgggtattgg ttgaatattt gtaaaccatt ccctagccta catatttatt actgaattaa    1080
ctttcctgat aaccattgca taattacatt tttctataaa atgaaagctt               1130
```

DERWENT-ACC-NO: 2000-572184

DERWENT-WEEK: 200065

TITLE: Breast tumor antigen polypeptides and polynucleotides, useful for manufacturing vaccines and compositions for treating, diagnosing, and monitoring breast cancer

INVENTOR: LODES, M J

PATENT-ASSIGNEE: CORIXA CORP[CORIN]

PRIORITY-DATA: 1999US-0396313 (September 17, 1999) , 1999US-0262505 (March 4, 1999) , 1999US-0272886 (March 19, 1999)

PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
WO 200052165 A2	September 8, 2000	E	140	C12N 015/12
AU 200033912 A	September 21, 2000	N/A	000	C12N 015/12

DESIGNATED-STATES: AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CR CU CZ DE DK DM EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT T Z UA UG US UZ VN YU ZA ZW AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL OA PT SD SE SL SZ TZ UG ZW

APPLICATION-DATA:

PUB-NO	APPL-DESCRIPTOR	APPL-NO	APPL-DATE
WO 200052165A2	N/A	2000WO-US05431	February 29, 2000
AU 200033912A	N/A	2000AU-0033912	February 29, 2000
AU 200033912A	Based on	WO 200052165	N/A

INT-CL (IPC): A61K031/7088, A61K038/17 , A61K039/395 , A61P035/00 , C07K014/82 , C07K016/30 , C12N005/08 , C12N015/12 , C12N015/62 , C12N015/63 , C12Q001/68 , G01N033/574

ABSTRACTED-PUB-NO: WO 200052165A

BASIC-ABSTRACT:

NOVELTY - A polypeptide comprising an immunogenic portion of a breast tumor antigen (I) or its variant, is new. The variant differs in one or more substitutions, deletions, additions and/or insertions, but its ability to react with antigen-specific antisera is not diminished.

DETAILED DESCRIPTION - A polypeptide comprising an immunogenic portion of a breast tumor antigen (I) or its variant, is new. The variant differs in one or more substitutions, deletions, additions and/or insertions, but its ability to react with antigen-specific antisera is not diminished.

(I) comprises an amino acid sequence that is encoded by a polynucleotide sequence consisting of:

(a) 154 fully defined polynucleotide sequences (derived from Homo sapiens) given in the specification,

consisting of 57 sequences having 38 base pairs (bp) (e. g. cagtagctagcatgcggacgactactactacgacgacg), or sequences having defined 266, 340, 422, 425, 426, 427, 431, 441, 448, 447, 451, 474, 522, 530, 535, 536, 540, 541, 551, 559, 577, 631, 643 or 651 bp (all given in the specification); and

(b) complements of (a).

INDEPENDENT CLAIMS are also included for the following:

(1) a polynucleotide encoding (I);

(2) expression vectors comprising the polynucleotide or its complement;

(3) a host cell transformed or transfected with the expression vector;

(4) an isolated antibody or its antigen-binding fragment that specifically binds to a (I);

(5) an antigen presenting cell that expresses the polypeptide;

(6) a fusion protein comprising at least 1 breast tumor antigen polypeptide;

(7) a polynucleotide encoding the fusion protein;

(8) vaccines or pharmaceutical compositions comprising the fusion protein, the polynucleotide, the polypeptide, the antibody or its antigen-binding fragment, or the antigen presenting cell;

(9) a method (M1) for removing tumor cells from a biological sample comprising contacting with T cells that specifically react with (I) to remove the cells that express the antigen;

(10) a method (M2) for stimulating and/or expanding T

cells specific for (I) comprising contacting the T cells with one or more of the polypeptide, the polynucleotide encoding (I), and/or an antigen presenting cell that expresses the polypeptide;

(11) an isolated T cell population prepared by M2;

(12) CD4+ and/or CD8+ T cells isolated from a patient and incubated with one or more of the polypeptides, the polynucleotide encoding it, and/or an antigen presenting cell that expresses the polypeptide such that the T cells proliferate;

(13) a method (M3) for inhibiting the development of breast cancer in a patient comprising:

(a) incubating CD4+ and/or CD8+ T cells isolated from a patient with one or more of the polypeptide, the polynucleotide encoding it, and/or an antigen presenting cell that expresses the polypeptide such that the T cells proliferate;

(b) cloning one or more proliferated cells; and

(c) administering the cloned T cells to the patient;

(14) methods (M4) for determining the presence or absence of a cancer in a patient;

(15) methods (M5) for monitoring the progression of cancer in a patient (no details given);

(16) diagnostic kits comprising antibodies and a detection reagent comprising a reporter group, or an oligonucleotide and a diagnostic reagent for use in a polymerase chain reaction (PCR) or hybridization assay; and

(17) an oligonucleotide comprising 10-40 nucleotides that hybridize under moderately stringent conditions to the polynucleotide that encodes (I).

ACTIVITY - Cytostatic.

No biological data given.

MECHANISM OF ACTION - Vaccine.

No biological data given.

USE - The polypeptide having an immunogenic portion of (I), the polynucleotide encoding the polypeptide, the antibody, or its antigen-binding fragment and the antigen presenting cell are useful for inhibiting the development of breast cancer in a patient. The isolated T cell population and the pharmaceutical compositions are also useful for inhibiting the development of breast cancer in a patient (all claimed). The breast tumor antigen polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for treating breast cancer, and for diagnosing and monitoring the cancer.

CHOSEN-DRAWING: Dwg.0/2

TITLE-TERMS: BREAST ANTIGEN USEFUL MANUFACTURE VACCINE COMPOSITION TREAT DIAGNOSE MONITOR BREAST CANCER

DERWENT-CLASS: B04 D16 S03

CPI-CODES: B04-B03C; B04-B04C2; B04-E03; B04-E05; B04-E08; B04-F0100E; B04-G01; B04-N0200E; B11-C07A; B11-C08E3; B11-C08E5; B12-K04A1; B12-K04F; B14-H01; B14-S11C; D05-H07; D05-H09; D05-H10; D05-H11; D05-H12A; D05-H12D1; D05-H12E; D05-H14; D05-H17A;

EPI-CODES: S03-E14H4;

CHEMICAL-CODES: Chemical Indexing M1 *01* Fragmentation Code M421 M423 M710 M905 N135 P633 Q233 Specfic Compounds A00H3T A00H3N

Chemical Indexing M1 *02* Fragmentation Code M421 M423 M710 M905 N135 P633 Q233 Specfic Compounds A00NST A00NSN

Chemical Indexing M1 *03* Fragmentation Code M421 M423 M710 M905 N135 P633 Q233 Specfic Compounds A012PT A012PN

Chemical Indexing M1 *04* Fragmentation Code M421 M423 M710 M905 N135 P633 Q233 Specfic Compounds A00C8T A00C8N

Chemical Indexing M1 *05* Fragmentation Code M423 M710 M750 M905 N135 P633 Q233 Specfic Compounds A00GTT A00GTN A00GTX

Chemical Indexing M1 *06* Fragmentation Code M423 M430 M782 M905 N102 N135 P831 Q233 Q505 Specfic Compounds A00C8K A00C8D A00C8M

Chemical Indexing M1 *07* Fragmentation Code M423 M430 M782 M905 N102 N135 P831 Q233 Q505 Specfic Compounds A013IK A013ID A013IM

Chemical Indexing M6 *08* Fragmentation Code M905 P633 P831 Q233 Q505 R515 R521 R621 R627 R639

SECONDARY-ACC-NO:

CPI Secondary Accession Numbers: C2000-170637

Non-CPI Secondary Accession Numbers: N2000-423264